

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 17:41:40 ; Search time 442 Seconds
(without alignments)
294.648 Million cell updates/sec

Title: US-10-848-737-1

Perfect score: 22
Sequence: 1 gugaacucacugagucutt 22

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 300 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002as:*
- 7: Geneseq2002bs:*
- 8: Geneseq2003as:*
- 9: Geneseq2003bs:*
- 10: Geneseq2003cs:*
- 11: Geneseq2003ds:*
- 12: Geneseq2004as:*
- 13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	90.9	29751	12	ADJ39000 SARS coro
2	16	72.7	619	10	ADK58384 Plant DNA
3	16	72.7	619	10	ADK57664 Plant DNA
4	16	72.7	619	11	ADMA45449 Insect re
5	16	72.7	672	11	ADMA44886 Insect re
6	16	72.7	686	10	ADK57646 Plant DNA
7	16	72.7	686	10	ADK54332 Plant DNA
8	16	72.7	696	11	ADMA45452 Insect re
9	15	68.2	534	6	ABN69216 Streptoco
10	15	68.2	537	8	ACA50276 Prokaryot
11	15	68.2	627	6	ABN61473 Prostata
12	15	68.2	1101	13	ADR61500 Cotton cd
13	15	68.2	1191	5	ABA14473 Human ner
14	15	68.2	1191	5	ABA14472 Human ner
15	15	68.2	1820	4	AAI22168 Probe #12
16	15	68.2	1820	4	ABA67247 Human foe
17	15	68.2	1820	4	ABA67247 Human foe
18	15	68.2	1820	4	AAI47467 Probe #16
19	15	68.2	1820	4	ABA49335 Human bre
20	15	68.2	1820	4	ABA34342 Probe #12
					AAK15686 Human bra

15	68.2	1820	5	AAI07869	Rai07869 Probe #78
15	68.2	1820	6	ABS15429	AbS15429 Human gen
15	68.2	1820	4	AAI24124	Aai24124 Probe #14
15	68.2	1820	4	ABA69247	AbA69247 Human foe
15	68.2	1820	4	AAI49420	Aai49420 Probe #18
15	68.2	1820	4	ABA51239	AbA51239 Human bre
15	68.2	1820	4	ABA36163	AbA36163 Probe #14
15	68.2	1820	4	AAK43349	AaK43349 Human bon
15	68.2	1820	4	AAK43349	AaK43349 Human bon
15	68.2	1820	4	AAK17541	AaK17541 Human bra
15	68.2	1820	6	ABS17430	AbS17430 Human gen
15	68.2	1820	4	AAI12930	Aai12930 Probe #28
15	68.2	1820	4	ABA54631	AbA54631 Human foe
15	68.2	1820	4	AAI34291	Aai34291 Probe #29
15	68.2	1820	4	ABA44181	AbA44181 Human bre
15	68.2	1820	4	ABA24415	AbA24415 Probe #28
15	68.2	1820	4	AAK02920	AaK02920 Human bra
15	68.2	1820	5	AAI02849	Aai02849 Probe #28
15	68.2	1820	6	ABS02874	AbS02874 Human gen
15	68.2	2394	10	ADC37626	AdC37626 Human nuc
15	68.2	4064	4	ABL01842	AbL01842 Drosophil
15	68.2	4064	4	ABL03058	AbL03058 Drosophil
15	68.2	4089	4	ABL04554	AbL04554 Drosophil
15	68.2	4317	13	ADR14707	AdR14707 Human NF-
15	68.2	43599	6	ABK84242	AbK84242 Human cDN
14	63.6	325	2	AAI19998	Aai19998 Human gen
14	63.6	348	6	ABN92664	AbN92664 Staphyloc
14	63.6	348	13	ADS01314	AdS01314 Staphyloc
14	63.6	534	6	ABN69215	AbN69215 Streptoco
14	63.6	1240	3	AAC53227	AdC53227 Arabidops
14	63.6	1278	13	ADT45310	AdT45310 Bacterial
14	63.6	1341	4	AAF28159	Aaf28159 Rice cent
14	63.6	1860	13	ADS16656	AdS16656 B. hensel
14	63.6	2350	10	ADE55033	AdE55033 Rat gene
14	63.6	2350	10	ADE55037	AdE55037 Rat gene
14	63.6	2457	10	ACF70959	AcF70959 Photorhab
14	63.6	3170	4	AAH54485	AaH54485 S. epider
14	63.6	3504	10	ADB69029	AdB69029 C. neofor
14	63.6	4314	8	ACA46888	AcA46888 Prokaryot
14	63.6	4323	5	AAH79450	AaH79450 DNA encod
14	63.6	4329	4	AAH53329	AaH53329 S. epider
14	63.6	9342	4	ABL02454	AbL02454 Drosophil
14	63.6	1895	13	ADS16655	AdS16655 B. hensel
14	63.6	47493	9	ADA02549	AdA02549 Mouse wnt
14	63.6	47493	10	ADB72287	AdB72287 Mouse wnt
14	63.6	47493	10	ADB95797	AdB95797 Mouse wnt
14	63.6	76180	13	ABD33385	AbD33385 Human can
14	63.6	76698	12	ADJ84182	AdJ84182 Human c-r
14	63.6	110000	4	AAI39682_07	Continuation (8 of
14	63.6	110000	4	AAI39683_07	Continuation (8 of
14	63.6	110000	6	ABN71527_00	Continuation (41 o
14	63.6	110000	10	ACF67367_40	Continuation (8 of
14	63.6	110000	10	ACF65388_07	Continuation (8 of
14	63.6	190000	10	ADL13752	AdL13752 Osteoarth
14	63.6	247544	12	ADQ59419	AdQ59419 Human can
13	59.1	21	6	ABA04625	AbA04625 MOLJ reve
13	59.1	21	10	ADJ87878	AdJ87878 G-coupled
13	59.1	21	12	ADN38524	AdN38524 Novel hum
13	59.1	190	2	AAQ76533	AaQ76533 Human gen
13	59.1	200	12	ACH89259	AcH89259 Human gen
13	59.1	240	11	ACH94421	AcH94421 Klebsiell
13	59.1	304	10	ADC72386	AdC72386 DNA Seq I
13	59.1	305	6	ABV88798	AbV88798 Human col
13	59.1	319	6	ABN19496	AbN19496 Human ORF
13	59.1	413	12	ADP95241	AdP95241 Cotton ex
13	59.1	438	3	AAH82241	AaH82241 N. mening
13	59.1	445	6	ABV97902	AbV97902 Human pan
13	59.1	478	4	ABA57952	AbA57952 Human foe
13	59.1	478	4	AAI37549	Aai37549 Probe #62
13	59.1	478	4	AAK31673	AaK31673 Human bon
13	59.1	478	4	AAK06027	AaK06027 Human bra
13	59.1	478	4	ABK31357	AbK31357 Human liv
13	59.1	478	6	ABS06429	AbS06429 Human gen
13	59.1	486	9	ACH20490	AcH20490 Human adu

C 94	13	59.1	539	4	AAK91848	Aak91848 Human cDN	167	13	59.1	2077	2	AAV90767	Av90767 Nucleotid
C 95	13	59.1	539	12	AAK93253	Aak93253 Human cDN	168	13	59.1	2105	8	ABT20603	Abt20603 Aspergill
C 96	13	59.1	539	12	ADL28275	Adl28275 5' end of	169	13	59.1	2105	8	ABT18787	Abt18787 Aspergill
C 97	13	59.1	539	12	ADL29680	Adl29680 5' end of	170	13	59.1	2160	13	ADR03843	Adr03843 Carotenoi
C 98	13	59.1	570	12	ACH75524	Ach75524 Human gen	171	13	59.1	2160	13	ADR03923	Adr03923 Method of
C 99	13	59.1	584	3	AAC35392	Aac35392 Arabidops	c 172	13	59.1	2418	6	ABZ13067	Abz13067 Arabidops
100	13	59.1	585	6	ABZ13833	Abz13833 Arabidops	c 173	13	59.1	2418	8	ADA68438	Ada68438 Arabidops
101	13	59.1	603	10	ADK60004	Adk60004 Plant DNA	c 174	13	59.1	2466	2	AAV90893	Av90893 Nucleotid
C 102	13	59.1	625	6	ABV96002	Abv96002 Human pan	c 175	13	59.1	2518	4	AAK94748	Aak94748 Human ful
C 103	13	59.1	646	6	ABV99049	Abv99049 Human pan	c 176	13	59.1	2518	12	ADL31785	Adl31785 Full leng
104	13	59.1	678	4	AAV53311	Aas53311 Haemophil	c 177	13	59.1	2555	12	ADQ23265	Adq23265 Human sof
105	13	59.1	678	8	ACA34154	Aca34154 Prokaryot	c 178	13	59.1	2571	13	ADR07375	Adr07375 Full leng
106	13	59.1	709	10	ADK56318	Adk56318 Plant DNA	c 179	13	59.1	2795	2	AAT97609	Aat97609 Human E2A
107	13	59.1	714	10	ADQ92613	Adq92613 Human col	c 180	13	59.1	2892	8	ACA24242	Aca24242 Prokaryot
C 108	13	59.1	796	6	ABN99064	Abn99064 Arabidops	c 181	13	59.1	2916	13	ADR85477	Adr85477 Aspergill
109	13	59.1	797	3	AAV78471	Aaa78471 Plant SDF	c 182	13	59.1	2933	5	AAH89905	Aah89905 Human bon
110	13	59.1	816	2	AAK30465	Aax30465 H. pylori	c 183	13	59.1	2991	4	ABK43659	Abk43659 DNA encod
111	13	59.1	816	2	AAK30429	Aax30429 H. pylori	c 184	13	59.1	2991	12	ADL54046	Adl54046 cDNA enco
112	13	59.1	816	4	AAS53726	Aas53726 Helicobac	c 185	13	59.1	3085	13	ADR84890	Adr84890 Aspergill
113	13	59.1	816	4	AAS53890	Aas53890 Helicobac	c 186	13	59.1	3170	6	ABZ35234	Abz35234 Human gen
114	13	59.1	816	8	ACA34980	Aca34980 Prokaryot	c 187	13	59.1	3273	10	ADF79065	Adf79065 Human pro
C 115	13	59.1	833	10	ACC43927	Acc43927 Nucleotid	c 188	13	59.1	3523	6	ABQ54610	Abq54610 Human ova
116	13	59.1	840	2	AAT68162	Aat68162 H. pylori	c 189	13	59.1	3849	4	AAI63871	Aai63871 Human pol
C 117	13	59.1	885	5	ABV30274	Abv30274 Human pro	c 190	13	59.1	3849	12	ADM24422	Adm24422 Human PRO
C 118	13	59.1	885	3	AAC77620	Aac77620 Human can	c 191	13	59.1	3854	2	AAT97610	Aat97610 Human E2A
119	13	59.1	944	6	ABK65152	Abk65152 Arabidops	c 192	13	59.1	3854	8	ABT18210	Abt18210 Aspergill
120	13	59.1	947	10	ADQ46634	Adq46634 Thalecres	c 193	13	59.1	3867	8	ABT20024	Abt20024 Aspergill
121	13	59.1	947	10	ADQ30295	Adq30295 Plant yle	c 194	13	59.1	3872	10	ADF79063	Adf79063 Human pro
122	13	59.1	947	12	ADL41550	Adl41550 Plant tra	c 195	13	59.1	3914	12	ADQ86121	Adq86121 Human tum
123	13	59.1	947	12	ADQ001654	Ado01654 Thalecres	c 196	13	59.1	3914	13	ACN38620	Acn38620 Tumour-as
124	13	59.1	954	5	AAH65497	Aah65497 C glutami	c 197	13	59.1	3916	6	ABZ78236	Abz78236 A. niger
125	13	59.1	1024	6	ABX65613	Abx65613 Helicobac	c 198	13	59.1	3919	13	ADT05424	Adt05424 Haemophil
C 126	13	59.1	1026	4	AAI81237	Aai81237 Human pol	c 199	13	59.1	3922	9	ACH03928	Ach03928 Human cDN
C 127	13	59.1	1065	5	AAV70168	Aav70168 DNA encod	c 200	13	59.1	3935	8	ABZ34770	Abz34770 Coding se
128	13	59.1	1137	13	ADS49156	Ades49156 Bacterial	c 201	13	59.1	3935	8	ACS50076	Acc50076 Breast ca
C 129	13	59.1	1140	3	AAC59746	Aac59746 Human sec	c 202	13	59.1	3935	10	ADF90853	Adf90853 Human hep
130	13	59.1	1146	8	ACA44595	Aca44595 Prokaryot	c 203	13	59.1	4025	10	ADE79067	Ade79067 Human pro
131	13	59.1	1161	10	ADQ30376	Adq30376 Bacterial	c 204	13	59.1	4033	13	ADR85839	Adr85839 Aspergill
132	13	59.1	1387	10	ADQ30625	Adq30625 Plant yle	c 205	13	59.1	4105	8	ABT18193	Abt18193 Aspergill
133	13	59.1	1387	12	ADL41852	Adl41852 Plant tra	c 206	13	59.1	4105	8	ABT20007	Abt20007 Aspergill
134	13	59.1	1387	12	ADQ62988	Ado62988 Transcrip	c 207	13	59.1	4125	13	ADR44022	Adr44022 Human col
135	13	59.1	1411	6	ABN97326	Abn97326 Gene #382	c 208	13	59.1	4657	13	ADR85252	Adr85252 Aspergill
136	13	59.1	1411	13	ADQ86272	Adq86272 Human tum	c 209	13	59.1	4894	6	ABA04591	Abaa04591 MOL3 codi
137	13	59.1	1411	13	ADQ87432	Adq87432 Human tum	c 210	13	59.1	4894	10	ADJ87653	Adj87653 G-coupled
138	13	59.1	1419	12	ADQ95933	Adq95933 T cell ac	c 211	13	59.1	5195	10	ADF59868	Adf59868 Human con
139	13	59.1	1422	12	ADQ95991	Adq95991 T cell ac	c 212	13	59.1	5733	3	AAA95416	Aaa95416 Aspergill
C 140	13	59.1	1476	4	AAI60351	Aai60351 Human pol	c 213	13	59.1	5869	10	ADE62578	Ades62578 Human gen
C 141	13	59.1	1539	10	ADC21639	Adc21639 Thogoto v	c 214	13	59.1	6373	4	ABLI0476	Abli0476 Drosophil
142	13	59.1	1546	4	AAI58565	Aai58565 Human pol	c 215	13	59.1	6657	13	ADR84665	Adr84665 Aspergill
143	13	59.1	1546	5	ADQ98782	Adq98782 DNA encod	c 216	13	59.1	7581	4	AAF85096	Aaf85096 Nucleotid
144	13	59.1	1546	9	ADB48542	Adb48542 Novel hum	c 217	13	59.1	7811	4	ABLI10750	Abli10750 Drosophil
C 145	13	59.1	1556	4	AAI63946	Aai63946 Human pol	c 218	13	59.1	9085	13	ADR84303	Adr84303 Aspergill
C 146	13	59.1	1556	4	AAK31613	Aak31613 cDNA enco	c 219	13	59.1	10303	4	ABLI13738	Abli13738 Drosophil
C 147	13	59.1	1556	4	ABK43947	Abk43947 DNA encod	c 220	13	59.1	11611	13	ADR03846	Adr03846 Carotenoi
C 148	13	59.1	1556	12	ADL54334	Adl54334 cDNA enco	c 221	13	59.1	11611	13	ADR03926	Adr03926 Genetical
C 149	13	59.1	1556	12	ADM24497	Adm24497 Human PRO	c 222	13	59.1	13160	5	ABA15804	Abai5804 Human ner
C 150	13	59.1	1583	3	AAC40650	Aac40650 Arabidops	c 223	13	59.1	13160	5	ADM20183	Adm20183 Alternati
C 151	13	59.1	1639	8	ADA73069	Ada73069 Rice gene	c 224	13	59.1	13189	5	ABA15805	Abai5805 Human ner
C 152	13	59.1	1683	10	ADE81196	Ades1196 Orf12 cod	c 225	13	59.1	13189	5	ADM20184	Adm20184 Alternati
153	13	59.1	1739	12	ADL42509	Adl42509 Plant tra	c 226	13	59.1	15739	13	ADR03845	Adr03845 Carotenoi
154	13	59.1	1767	8	ABT21218	Abt21218 Aspergill	c 227	13	59.1	15739	13	ADR03925	Adr03925 Genetical
155	13	59.1	1775	12	ADL42510	Adl42510 plant tra	c 228	13	59.1	16103	13	ADR03904	Adr03904 Carotenoi
C 156	13	59.1	1778	10	ADB53775	Adb53775 Primary r	c 229	13	59.1	16103	13	ADR03984	Adr03984 Genetical
C 157	13	59.1	1778	12	ADQ24857	Adq24857 Rat novel	c 230	13	59.1	16161	6	AAD36071	Aad36071 Human adi
158	13	59.1	1791	8	ABT19398	Abt19398 Aspergill	c 231	13	59.1	16245	13	ADR03878	Adr03878 B trispor
C 159	13	59.1	1830	6	ABK94898	Abk94898 Human nov	c 232	13	59.1	16245	13	ADR03958	Adr03958 Genetical
160	13	59.1	1854	8	ABT18804	Abt18804 Aspergill	c 233	13	59.1	16954	13	ADR03886	Adr03886 Carotenoi
161	13	59.1	1867	8	ABT20620	Abt20620 Aspergill	c 234	13	59.1	16954	13	ADR03885	Adr03885 Carotenoi
162	13	59.1	2000	8	ADA73355	Ada73355 Rice Gene	c 235	13	59.1	16954	13	ADR03965	Adr03965 Genetical
163	13	59.1	2025	8	ABT19381	Abt19381 Aspergill	c 236	13	59.1	16954	13	ADR03966	Adr03966 Genetical
164	13	59.1	2025	8	ABT21201	Abt21201 Aspergill	c 237	13	59.1	17118	13	ADR03890	Adr03890 Carotenoi
C 165	13	59.1	2047	6	ABL99809	AbL99809 Human sec	c 238	13	59.1	17118	13	ADR03970	Adr03970 Genetical
C 166	13	59.1	2047	8	ABX62972	Abx62972 Human act	c 239	13	59.1	17238	13	ADR03881	Adr03881 Carotenoi

Qy 6 CUCACUCGUGAGCUCT 21

```

Db      622 CTCACCTGAGCTCT 637
      |||:|||||:|||||:
RESULT 9
ABN69216
ID ABN69216 standard; DNA; 534 BP.
XX AC ABN69216;
XX DT 01-JUL-2002 (first entry)
XX DE Streptococcus polynucleotide SEQ ID NO 6345.
XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
XX KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX OS Streptococcus pyogenes.
XX PN WO200234771-A2.
XX PD 02-MAY-2002.
XX PF 29-OCT-2001; 2001WO-GB004789.
XX PR 27-OCT-2000; 2000GB-00026333.
XX PR 24-NOV-2000; 2000GB-00028727.
XX PR 07-MAR-2001; 2001GB-00005640.
XX PA (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RES.
XX PI Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C;
XX PI Tettelin H;
XX DR WPI; 2002-352536/38.
XX DR P-PSDB; ABP28585.
XX PT New Streptococcus protein for the treatment or prevention of infection or
XX PT disease caused by Streptococcus bacteria, such as meningitis, and for
XX PT detecting a compound that binds to the protein.
XX PS Claim 7; Page 3799; 4525pp; English.
XX CC The invention relates to a protein (ABP25413-ABP30895) from group B
XX CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
XX CC the specification. The proteins have antibacterial and antiinflammatory
XX CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
XX CC antibodies that bind (I) are used in the manufacture of medicaments for
XX CC the treatment or prevention of infection or disease caused by
XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX CC Nucleic acids encoding (I) are used to detect Streptococcus in a
XX CC biological sample. (I) is used to determine whether a compound binds to
XX CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
XX CC used as a vaccine or diagnostic composition. The disease caused by
XX CC Streptococcus that is prevented or treated may be meningitis. Nucleic
XX CC acid encoding (I) may be used to recombinantly produce (I) and may be
XX CC used in gene therapy. Antibodies to (I) are used for affinity
XX CC chromatography, immunoassays, and distinguishing/identifying
XX CC Streptococcus proteins
XX SQ Sequence 534 BP; 186 A; 101 C; 109 G; 138 T; 0 U; 0 Other;
Query March 68.2%; Score 15; DB 6; Length 534;
Best Local Similarity 80.0%; Pred. NO. 29;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Oy 3 GAACUCACUCGAG 17
      |||:|||||:|||||:
Db      94 GAACCTACTGAG 108
      |||:|||||:|||||:
RESULT 10
ACA50276
ID ACA50276 standard; DNA; 537 BP.
XX AC ACA50276;
XX DT 19-JUN-2003 (first entry)
XX DE Prokaryotic essential gene #31933.
XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;
XX KW drug design; gene.
XX OS Streptococcus pyogenes.
XX PN WO200277183-A2.
XX PD 03-OCT-2002.
XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 06-SEP-2001; 2001US-00815242.
XX PR 25-OCT-2001; 2001US-00948993.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR P-PSDB; ABU46406.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 14; SEQ ID NO 38146; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
XX CC prokaryotic essential genes. Note: The sequence data for this patent did
XX CC not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences

```

```
XX SQ Sequence 537 BP; 188 A; 101 C; 109 G; 139 T; 0 U; 0 Other;
Query Match 68.2%; Score 15; DB 8; Length 537;
Best Local Similarity 80.0%; Pred. No. 29;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 GAACUCACUCUGAG 17
||||:||||:||||
DB 94 GAACCTCACTCGTGAG 108

RESULT 11
ABS61473/c
ID ABS61473 standard; cDNA; 627 BP.
XX AC
XX ABS61473;
XX
XX 05-NOV-2002 (first entry)
XX
XX Prostate specific nucleic acid DEX0259_81.
XX
XX Human; ss; prostate specific nucleic acid; PSNA; prostate cancer;
XX cytostatic; non-cancerous prostate disease; PSP; vaccine;
XX prostate specific protein; metastasis.
XX
XX Homo sapiens.
XX
XX WO200242776-A2.
XX
XX 30-MAY-2002.
XX
XX 01-NOV-2001; 2001WO-US045654.
XX
XX 01-NOV-2000; 2000US-0244782P.
XX
XX (DIAD-) DIADEXUS INC.
XX
XX Sun Y, Recipon H, Chen S, Liu C;
XX
XX WPI; 2002-490217/52.
XX
XX New polypeptide useful for diagnosing and monitoring the presence and
XX metastases of prostate cancer in a patient and as a component in
XX databases for search analysis as well as in sequence analysis algorithms.
XX
XX Claim 1; Page 183-184; 242pp; English.
XX
XX The invention relates to an isolated polypeptide comprising a sequence
XX with 60 % identity to one of prostate specific protein (PSP) sequences,
XX or comprising an amino acid sequence encoded by one of 136 nucleotide
XX prostate specific nucleic acid sequences, PSNA, (or a sequence that
XX hybridises to it or is 60% identical to it), given in the specification.
XX Also included are a vector comprising the polynucleotide, a host cell
XX comprising the vector, an antibody specific for the PSP proteins and a
XX vaccine comprising the protein or polynucleotide. The PSP and PSNA are
XX useful for diagnosing and monitoring the presence and metastases of
XX prostate cancer in a patient. The PSNA is useful for determining the
XX level PSNA in a sample. An antibody to the PSP is useful for determining
XX the presence of prostate specific protein in a sample, and for treating a
XX patient with prostate cancer, which induces an immune response against
XX the prostate cancer cell expressing the nucleic acid or polypeptide and a
XX kit is useful for detecting a risk of cancer or presence of cancer in a
XX patient. PSNA is useful as hybridisation probes to detect, characterise
XX and quantify hybridising nucleic acids from both genomic and transcript-
XX derived nucleic acid samples and also in microarrays. Sequences of PSP
XX and PSNA are useful as components in databases for search analysis as
XX well as in sequence analysis algorithms. PSNA is useful to drive in vivo
XX expression of PSP. The present sequence is a PSNA of the invention
XX
XX SQ Sequence 627 BP; 129 A; 152 C; 169 G; 177 T; 0 U; 0 Other;
Query Match 68.2%; Score 15; DB 6; Length 627;

Best Local Similarity 80.0%; Pred. No. 29;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 GAACUCACUCUGAG 17
||||:||||:||||
DB 224 GAACCTCACTCGTGAG 210

RESULT 12
ADR61500
ID ADR61500 standard; cDNA; 1101 BP.
XX AC
XX ADR61500;
XX
XX 02-DEC-2004 (first entry)
XX
XX Cotton cDNA sequence, SEQ ID 2281.
XX
XX Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;
XX drought tolerance; plant disease resistance; galactomannan; lignin;
XX plant growth regulator; heat tolerance; herbicide tolerance;
XX homologous recombination; extreme osmotic condition tolerance;
XX pathogen resistance; pest resistance; yield; photosynthesis; seed oil;
XX stress resistance.
XX
XX Gossypium hirsutum.
XX
XX US2004181830-A1.
XX
XX 16-SEP-2004.
XX
XX 29-JAN-2004; 2004US-00767795.
XX
XX 07-MAY-2001; 2001US-00849529.
XX
XX 12-DEC-2001; 2001US-00021323.
XX
XX (KOVA/) KOVALIC D K.
XX (ZHOU/) ZHOU Y.
XX (CAOY/) CAO Y.
XX
XX Kovalic DK, Zhou Y, Cao Y;
XX
XX WPI; 2004-667718/65.
XX
XX New recombinant nucleic acid molecules and polypeptides from Gossypium
XX hirsutum, useful for producing plants with improved biological
XX characteristics (e.g. improved plant cold or drought tolerance).
XX
XX Claim 1; SEQ ID NO 2281; 14pp; English.
XX
XX The invention relates to a recombinant polynucleotide comprising any of
XX the 58798 Cotton plant cDNA sequences mentioned in the specification.
XX Also a recombinant polypeptide comprising any of the 58798 amino acid
XX sequences mentioned in the specification and producing a plant having an
XX improved property. Producing a plant having an improved property
XX comprises transforming a plant with a recombinant construct comprising a
XX promoter region functional in a plant cell operably joined to a
XX polynucleotide comprising a coding sequence for a polypeptide associated
XX with the property, and growing the transformed plant. The polypeptide is
XX useful for improving plant cold tolerance, manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, improving plant
XX drought tolerance, providing increased resistance to plant disease,
XX producing galactomannan (or lignin or plant growth regulators), improving
XX plant heat tolerance, improving plant tolerance to herbicides, increasing
XX the rate of homologous recombination in plants, improving plant tolerance
XX to extreme osmotic conditions or to pathogens or pests, improving yield
XX by modification of photosynthesis, modifying seed oil or protein yield
XX and/or content, improving yield by modification of carbohydrate, nitrogen
XX or phosphorus use and/or uptake, or improving yield by providing improved
XX plant growth and development under at least one stress condition. The
XX polynucleotide and polypeptide may also be used in recombinant DNA
XX constructs, in physical arrays of molecules, as plant breeding markers,
XX or in computer-based storage and analysis systems. The present sequence
```

CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040181830. However only 6585
CC polynucleotide sequences were available, the remaining 52213
CC polynucleotides and all 58798 protein sequences were not present.

XX SQ Sequence 1101 BP; 308 A; 201 C; 254 G; 338 T; 0 U; 0 Other;

Query Match 68.2%; Score 15; DB 13; Length 1101;
Best Local Similarity 73.3%; Pred. No. 29;
Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CUCACUCUGAGTUC 20
|:|:|:|:|:|:|:
Db 1027 CTCACCTCGAGTTC 1041

RESULT 13

ABAI4473
ID ABA14473 standard; DNA; 1191 BP.

XX AC ABA14473;

XX DT 23-JAN-2002 (first entry)

XX DE Human nervous system related polynucleotide SEQ ID NO 6804.

XX KW Human; nontropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antickling; antianaemic; antithratic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antilucer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

XX OS Homo sapiens.

XX PN WO200159063-A2.

XX PD 16-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001334.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225213P.

XX PR 14-AUG-2000; 2000US-0225214P.

XX PR 14-AUG-2000; 2000US-0225266P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

XX PR 14-AUG-2000; 2000US-0225757P.

XX PR 14-AUG-2000; 2000US-0225758P.

XX PR 14-AUG-2000; 2000US-0225759P.

XX PR 14-AUG-2000; 2000US-0225760P.

PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244647P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.


```
XX OS Homo sapiens.
XX PN WO200157278-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000670.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PR WPI; 2001-488901/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human cervical epithelial cells.
XX PS Claim 25; SEQ ID NO 12101; 487pp; English.
XX CC The present invention relates to human single exon nucleic acid probes
XX CC (SENPs). The present sequence is one such probe. The SENPs are derived
XX CC from human HeLa cells. The SENPs can be used to produce a single exon
XX CC microarray, which can be used for measuring human gene expression in a
XX CC sample derived from human cervical epithelial cells. By measuring gene
XX CC expression, the probes are therefore useful in grading and/or staging of
XX CC diseases of the cervix, notably cervical cancer. Note: The sequence data
XX CC for this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1820 BP; 546 A; 406 C; 445 G; 423 T; 0 U; 0 Other;

Query Match 68.2%; Score 15; DB 4; Length 1820;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 CACUCGUGAGCUCTT 22
DB 1774 CACTCGTGAGCTCTT 1760

RESULT 16
ABA67247/c
ID ABA67247 standard; DNA; 1820 BP.
XX AC ABA67247;
XX DT 01-FEB-2002 (first entry)
XX DE Human foetal liver single exon nucleic acid probe #15552.
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX OS Homo sapiens.
XX PN WO200157277-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000669.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.

Query Match 68.2%; Score 15; DB 4; Length 1820;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 CACUCGUGAGCUCTT 22
DB 1774 CACTCGTGAGCTCTT 1760

RESULT 17
AAI47467/c
ID AAI47467 standard; DNA; 1820 BP.
XX AC AAI47467;
XX DT 17-OCT-2001 (first entry)
XX DE Probe #16153 used to measure gene expression in human placenta sample.
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000663.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PR WPI; 2001-488997/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human placenta.
```

```
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PR WPI; 2001-483447/52.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human fetal liver.
XX PS Claim 4; SEQ ID NO 15552; 639pp + Sequence Listing; English.
XX CC The invention relates to a single exon nucleic acid probe for measuring
XX CC human gene expression in a sample derived from human foetal liver. The
XX CC single exon nucleic acid probes may be used for predicting, measuring and
XX CC displaying gene expression in samples derived from human fetal liver. The
XX CC present sequence is a single exon nucleic acid probe of the invention.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1820 BP; 546 A; 406 C; 445 G; 423 T; 0 U; 0 Other;

Query Match 68.2%; Score 15; DB 4; Length 1820;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 CACUCGUGAGCUCTT 22
DB 1774 CACTCGTGAGCTCTT 1760

RESULT 17
AAI47467/c
ID AAI47467 standard; DNA; 1820 BP.
XX AC AAI47467;
XX DT 17-OCT-2001 (first entry)
XX DE Probe #16153 used to measure gene expression in human placenta sample.
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000663.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PR WPI; 2001-488997/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human placenta.
```

XX Claim 25; SEQ ID NO 16153; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 1820 BP; 546 A; 406 C; 445 G; 423 T; 0 U; 0 Other;
Query Match 68.2%; Score 15; DB 4; Length 1820;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 8 CACUCGUGAGCUCU 22
DB 1774 CACTCGTGAGCTCTT 1760
RESULT 18
ABA49335/c
ID ABA49335 standard; DNA; 1820 BP.
AC ABA49335;
XX
XX 01-FEB-2002 (first entry)
XX
XX Human breast cell single exon nucleic acid probe #8030.
XX Human; microarray; single exon probe; gene expression; breast; disease;
XX cancer; ss.
XX Homo sapiens.
XX
XX WO200157271-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000662.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes, useful
XX for measuring gene expression in sample derived from human breast,
XX comprises number of single exon nucleic acid probes.
XX
XX Claim 4; SEQ ID NO 8030; 327pp + Sequence Listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BT 474 cells. The method involves contacting the
XX probes with a collection of detectably labelled nucleic acids derived
XX from mRNA of human breast, and then measuring the label bound to each
XX probe of the microarray. The probes are useful for verifying the
XX expression of regions of genomic DNA predicted to encode proteins. They
XX are useful for gene discovery, and for determining predisposition and/or
XX prognosing breast disease. Gene expression analysis is useful for
XX assessing the toxicity of chemical agents on cells. The microarray of
XX this invention presents a far greater diversity of probes for measuring
XX gene expression, with far less bias than expressed sequence tag

CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1820 BP; 546 A; 406 C; 445 G; 423 T; 0 U; 0 Other;
Query Match 68.2%; Score 15; DB 4; Length 1820;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 8 CACUCGUGAGCUCU 22
DB 1774 CACTCGTGAGCTCTT 1760
RESULT 19
ABA34342/c
ID ABA34342 standard; DNA; 1820 BP.
XX
XX ABA34342;
XX
XX 23-JAN-2002 (first entry)
XX
XX Probe #12808 for gene expression analysis in human heart cell sample.
XX Human; gene expression; heart; microarray; vascular system; probe;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease; ss.
XX Homo sapiens.
XX
XX WO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000666.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts.
XX
XX Claim 4; SEQ ID NO 12808; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart. The
XX present sequence is one such probe. The probes may be used for
XX predicting, measuring and displaying gene expression in samples derived
XX from the human heart via microarrays. By measuring gene expression, the
XX probes are useful for predicting, diagnosing, grading, staging,
XX monitoring and prognosing diseases of the human heart and vascular system
XX e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX congenital heart disease. Note: The sequence data for this patent did not
XX form part of the printed specification, but was obtained in electronic
XX format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1820 BP; 546 A; 406 C; 445 G; 423 T; 0 U; 0 Other;

```

Query Match      68.2%; Score 15; DB 4; Length 1820;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      8 CACUCGUGAGCUCTT 22
Db      1774 CACTCGTGAGCTCTT 1760

RESULT 20
AAK15686/c
ID      AAK15686 standard; DNA; 1820 BP.
XX
AC      AAK15686;
XX
DT      05-NOV-2001 (first entry)
XX
DE      Human brain expressed single exon probe SEQ ID NO: 15677.
XX
KW      Human; brain expressed exon; gene expression analysis; probe; microarray;
KW      Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW      ss.
XX
OS      Homo sapiens.
XX
PN      W0200157275-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001WO-US000667.
XX
PR      04-FEB-2000; 2000US-0180312P.
PR      26-MAY-2000; 2000US-0207456P.
PR      30-JUN-2000; 2000US-00608408.
PR      03-AUG-2000; 2000US-00632366.
PR      21-SEP-2000; 2000US-0234687P.
PR      27-SEP-2000; 2000US-0236359P.
PR      04-OCT-2000; 2000GB-00024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-483446/52.
XX
Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS      Example 4; SEQ ID NO 15677; 650pp + Sequence Listing; English.
XX
CC      The present invention provides a number of single exon nucleic acid
CC      probes which are derived from genomic sequences expressed in the human
CC      brain. They can be used to measure gene expression in brain cell samples,
CC      which may enable the diagnosis and improved treatment of nervous system
CC      diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC      epilepsy and cancers. The present sequence is one of the probes of the
CC      invention
XX
SQ      Sequence 1820 BP; 546 A; 406 C; 445 G; 423 T; 0 U; 0 Other;

Query Match      68.2%; Score 15; DB 4; Length 1820;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      8 CACUCGUGAGCUCTT 22
Db      1774 CACTCGTGAGCTCTT 1760

RESULT 21
AAI07869/c
ID      AAI07869 standard; DNA; 1820 BP.
XX

```

```

AC      AAI07869;
XX
DT      09-OCT-2001 (first entry)
XX
DE      Probe #7860 used to measure gene expression in human breast sample.
XX
KW      Probe; human; breast disease; breast cancer; development disorder; ss;
KW      inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS      Homo sapiens.
XX
PN      W0200157270-A2.
XX
PD      09-AUG-2001.
XX
PF      29-JAN-2001; 2001WO-US000661.
XX
PR      04-FEB-2000; 2000US-0180312P.
PR      26-MAY-2000; 2000US-0207456P.
PR      30-JUN-2000; 2000US-00608408.
PR      03-AUG-2000; 2000US-00632366.
PR      21-SEP-2000; 2000US-0234687P.
PR      27-SEP-2000; 2000US-0236359P.
PR      04-OCT-2000; 2000GB-00024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-476286/51.
XX
Novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
XX
PS      Claim 25; SEQ ID NO 7860; 322pp; English.
XX
CC      The present invention relates to novel single exon nucleic acid probes.
CC      The present sequence is one such probe. The probes are useful for
CC      measuring human gene expression in a human breast sample, where the probe
CC      hybridises at high stringency to a nucleic acid expressed in the human
CC      breast. The probes are useful for predicting, diagnosing, grading,
CC      staging, monitoring and prognosing diseases of the human breast,
CC      particularly those diseases with polygenic aetiology. The diseases
CC      include: breast cancer, disorders of development, inflammatory diseases
CC      of the breast, fibrocystic changes, proliferative breast disease and non-
CC      carcinoma tumours. Note: The sequence data for this patent did not form
CC      part of the printed specification, but was obtained in electronic format
CC      directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ      Sequence 1820 BP; 546 A; 406 C; 445 G; 423 T; 0 U; 0 Other;

Query Match      68.2%; Score 15; DB 5; Length 1820;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      8 CACUCGUGAGCUCTT 22
Db      1774 CACTCGTGAGCTCTT 1760

RESULT 22
ABS15429/c
ID      ABS15429 standard; DNA; 1820 BP.
XX
AC      ABS15429;
XX
DT      19-AUG-2002 (first entry)
XX
DE      Human genome-derived single exon probe ORF from lung SEQ ID NO 15420.
XX
KW      Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW      chronic obstructive pulmonary disease; interstitial lung disease;
KW      familial idiopathic pulmonary fibrosis; neurofibromatosis;

```

KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
XX WO200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US0000665.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX
XX Claim 4; SEQ ID NO 15420; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of probes
XX ; the novel set of probes which hybridise at high stringency to a nucleic
XX acid expressed in the human lung; measuring gene expression in a sample
XX derived from human lung, comprising (a) contacting the array with a
XX collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of the
XX array; identifying exons in a eukaryotic genome, comprising (a)
XX algorithmically predicting at least one exon from genomic sequences of
XX the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarrays having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 12011 sequences, mentioned in the specification, or encoded by the
XX probes/open reading frames (ORF). The probes are used for gene expression
XX analysis, and for identifying exons in a gene, particularly using human
XX lung derived mRNA and for the study of lung diseases such as asthma, lung
XX cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
XX disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
XX tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
XX Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
XX histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
XX Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
XX dyskinesia, pulmonary hypertension and hyaline membrane disease. The
XX present sequence is a single exon probe open reading frame of the
XX invention. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 1820 BP; 546 A; 406 C; 445 G; 423 T; 0 U; 0 Other;
Query Match 68.2%; Score 15; DB 6; Length 1820;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 8 CACUCGUGAGCUCTT 22
Db 1774 CACTCGTGAGCTCTT 1760
|||:||||:||||
RESULT 23
AAI24124/C
ID AAI24124 standard; DNA; 1870 BP.
XX
XX AC AAI24124;
XX
XX DT 12-OCT-2001 (first entry)
XX
XX DE Probe #14057 for gene expression analysis in human cervical cell sample.
XX
XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX KW cervical cancer; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157278-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US0000670.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human cervical epithelial cells.
XX
XX Claim 25; SEQ ID NO 14057; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SENPs). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging of
XX diseases of the cervix, notably cervical cancer. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1870 BP; 568 A; 416 C; 455 G; 431 T; 0 U; 0 Other;
Query Match 68.2%; Score 15; DB 4; Length 1870;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 8 CACUCGUGAGCUCTT 22
Db 1774 CACTCGTGAGCTCTT 1760
|||:||||:||||

```
RESULT 24
ABA69247/c
ID ABA69247 standard; DNA; 1870 BP.
XX
XX
PD
XX
XX
AC ABA69247;
XX
XX
DT 01-FEB-2002 (first entry)
XX
XX
DE Human foetal liver single exon nucleic acid probe #17552.
XX
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200157277-A2.
XX
XX
PD 09-AUG-2001.
XX
XX
XX 30-JAN-2001; 2001WO-US000669.
XX
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX
DR WPI; 2001-488897/53.
XX
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
gene expression in human placenta.
XX
XX
PS Claim 4; SEQ ID NO 17552; 639pp + Sequence Listing; English.
XX
XX
CC The invention relates to a single exon nucleic acid probe for measuring
human gene expression in a sample derived from human foetal liver. The
single exon nucleic acid probes may be used for predicting, measuring and
displaying gene expression in samples derived from human foetal liver. The
present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 1870 BP; 568 A; 416 C; 455 G; 431 T; 0 U; 0 Other;

Query Match 68.2%; Score 15; DB 4; Length 1870;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 CACUCGUGAGCUCTT 22
DB 1774 CACTCGTGAGCTCTT 1760

RESULT 25
AAI49420/c
ID AAI49420 standard; DNA; 1870 BP.
XX
XX
AC AAI49420;
XX
XX
DT 17-OCT-2001 (first entry)
XX
XX
DE Probe #18106 used to measure gene expression in human placenta sample.
XX
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
genetic disorder; ss.
XX
XX
OS Homo sapiens.

Query Match 68.2%; Score 15; DB 4; Length 1870;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 CACUCGUGAGCUCTT 22
DB 1774 CACTCGTGAGCTCTT 1760

RESULT 26
ABA51239/c
ID ABA51239 standard; DNA; 1870 BP.
XX
XX
AC ABA51239;
XX
XX
DT 01-FEB-2002 (first entry)
XX
XX
DE Human breast cell single exon nucleic acid probe #9934.
XX
XX
KW Human; microarray; single exon probe; gene expression; breast; disease;
cancer; ss.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200157271-A2.
XX
XX
PD 09-AUG-2001.
XX
XX
XX 30-JAN-2001; 2001WO-US000662.
XX
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX
DR WPI; 2001-488897/53.
XX
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
gene expression in human placenta.
XX
XX
PS Claim 25; SEQ ID NO 18106; 654pp; English.
XX
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
The present sequence is one such probe. The probes are useful for
producing a microarray for predicting, measuring and displaying gene
expression in samples derived from human placenta. The probes are useful
for antenatal diagnosis of human genetic disorders
XX
XX
SQ Sequence 1870 BP; 568 A; 416 C; 455 G; 431 T; 0 U; 0 Other;

Query Match 68.2%; Score 15; DB 4; Length 1870;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 CACUCGUGAGCUCTT 22
DB 1774 CACTCGTGAGCTCTT 1760

RESULT 26
ABA51239/c
ID ABA51239 standard; DNA; 1870 BP.
XX
XX
AC ABA51239;
XX
XX
DT 01-FEB-2002 (first entry)
XX
XX
DE Human breast cell single exon nucleic acid probe #9934.
XX
XX
KW Human; microarray; single exon probe; gene expression; breast; disease;
cancer; ss.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200157271-A2.
XX
XX
PD 09-AUG-2001.
XX
XX
XX 30-JAN-2001; 2001WO-US000662.
XX
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
```

```
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
PS Claim 4; SEQ ID NO 9934; 327pp + Sequence Listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and Br 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1870 BP; 568 A; 416 C; 455 G; 431 T; 0 U; 0 Other;
Query Match 68.2%; Score 15; DB 4; Length 1870;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 8 CACUCGUGAGCUCTT 22
Db 1774 CACTCGTGAGCTCTT 1760
RESULT 27
ABA36163/c
ID ABA36163 standard; DNA; 1870 BP.
AC ABA36163;
XX
XX 23-JAN-2002 (first entry)
XX
XX Probe #14629 for gene expression analysis in human heart cell sample.
XX Human; gene expression; heart; microarray; vascular system; probe;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease; ss.
XX
XX Homo sapiens.
XX
XX WO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000666.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
PT
XX Example 4; SEQ ID NO 17906; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
```

```
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488999/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
PT
XX Claim 4; SEQ ID NO 14629; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1870 BP; 568 A; 416 C; 455 G; 431 T; 0 U; 0 Other;
Query Match 68.2%; Score 15; DB 4; Length 1870;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 8 CACUCGUGAGCUCTT 22
Db 1774 CACTCGTGAGCTCTT 1760
RESULT 28
AAK43349/c
ID AAK43349 standard; DNA; 1870 BP.
XX
XX AAK43349;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed single exon probe SEQ ID NO: 17906.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
XX Homo sapiens.
XX
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000668.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
PT
XX Example 4; SEQ ID NO 17906; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
```

CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention

XX
 SQ Sequence 1870 BP; 568 A; 416 C; 455 G; 431 T; 0 U; 0 Other;

Query Match 58.2%; Score 15; DB 4; Length 1870;
 Best Local Similarity 80.0%; Pred. No. 30;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 8 CACUCGUGAGCUCTT 22

Db 1774 CACTCGTGAGCTCTT 1760
 |||:|:|:|:|:|:|

RESULT 29

AAK17541/c

ID AAK17541 standard; DNA; 1870 BP.

XX AAK17541;

XX 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe SEQ ID NO: 17532.

XX Human; brain expressed exon; gene expression analysis; probe; microarray;
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
 KW 86.

XX Homo sapiens.

XX WO200157275-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000667.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.

XX Example 4; SEQ ID NO 17532; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention

XX Sequence 1870 BP; 568 A; 416 C; 455 G; 431 T; 0 U; 0 Other;

Query Match 58.2%; Score 15; DB 4; Length 1870;
 Best Local Similarity 80.0%; Pred. No. 30;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 8 CACUCGUGAGCUCTT 22

Db 1774 CACTCGTGAGCTCTT 1760
 |||:|:|:|:|:|:|

RESULT 30

ABS17430/c

ID ABS17430 standard; DNA; 1870 BP.

XX ABS17430;

XX 19-AUG-2002 (first entry)

XX Human genome-derived single exon probe ORF from lung SEQ ID NO 17421.

XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

KW chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;

KW hyaline membrane disease; open reading frame; ORF.

XX Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US000665.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.

XX Claim 4; SEQ ID NO 17421; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC; the novel set of probes which hybridise at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that

CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karagane syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1870 BP; 568 A; 416 C; 455 G; 431 T; 0 U; 0 Other;
Query Match 68.2%; Score 15; DB 6; Length 1870;
Best Local Similarity 80.0%; Pred. No. 30;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 8 CACUCGUGAGCTCTT 22
Db 1774 CACTCGTGAGCTCTT 1760
Search completed: July 20, 2005, 23:34:20
Job time : 455 secs

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 22:37:01 ; Search time 3090 Seconds
(without alignments)
271.008 Million cell updates/sec

Title: US-10-848-737-1

Perfect score: 22

Sequence: 1 gugaacacucgagacuttt 22

Scoring table:

OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 300 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hlc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gse1:*

9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	17	77.3	1344	4	BG917388
C 2	16	72.7	200	7	CF296001
C 3	16	72.7	206	7	CF29826
C 4	16	72.7	245	7	CF298930
C 5	16	72.7	270	2	BE228712
C 6	16	72.7	272	7	CF295034
C 7	16	72.7	273	7	CF299062
C 8	16	72.7	276	7	CF300491
C 9	16	72.7	277	7	CF300259
C 10	16	72.7	281	7	CF298981
C 11	16	72.7	283	7	CF298536
C 12	16	72.7	311	7	CF294911
C 13	16	72.7	313	6	CF299992
C 14	16	72.7	313	6	CB668085
C 15	16	72.7	321	7	CF300735
C 16	16	72.7	321	7	CF300267
C 17	16	72.7	321	7	CF334972
C 18	16	72.7	324	7	CF301775
C 19	16	72.7	326	7	CF295092
C 20	16	72.7	329	7	CF300479
C 21	16	72.7	329	7	CF298299
C 22	16	72.7	330	7	CF301705
C 23	16	72.7	340	7	CF299101
C 24	16	72.7	342	7	CF294601

CF300860	7LEAF--05
CF302308	7LEAF--07
BF10487	7009E03 O
CF333339	JMT--03-A
CF300157	7LEAF--04
CF301059	7LEAF--05
CF294542	30DGS--04
CF293314	30DGS--02
CF301046	7LEAF--06
CF301825	7LEAF--05
CF295099	30DGS--04
CF293177	30DGS--02
CF333477	JMT--02-G
CF294625	30DGS--04
BF797027	H062C08 E
CF294049	30DGS--03
CF300473	7LEAF--04
CF302540	7LEAF--08
CF300089	7LEAF--04
CF334002	JMT--03-C
CF293465	30DGS--02
CF294528	30DGS--04
CF294797	30DGS--04
CF334668	JMT--04-A
CF298619	7LEAF--02
CF299944	7LEAF--04
CF301512	7LEAF--06
CA882915	SSH19 Ri
CF333248	JMT--02-B
BE228607	98AS2865
CF296172	30DGS--06
CF302115	7LEAF--07
CF279865	14ETL--06
CF292628	30DGS--01
CF293499	30DGS--02
CF295674	30DGS--05
CF281031	14ETL--07
CF293653	30DGS--02
CF301091	7LEAF--05
CF292898	30DGS--01
CF301882	7LEAF--06
CF333799	JMT--02-N
CF293571	30DGS--02
CF292038	14ROOT--0
CF291160	14ROOT--0
CF295956	30DGS--06
CF297322	30DGS--08
CF302167	7LEAF--07
CF292634	30DGS--01
CF292393	30DGS--01
CF317419	HD--07-B1
CF302841	7LEAF--08
CF37043	JMT--07-F
CF310192	ABF--04-L
CA752917	JSYL446 r
CA752923	JSYL471 r
CA752899	JSYL374 r
CF293078	30DGS--02
CF301435	7LEAF--06
AM069961	MG1e1002A
CF293708	30DGS--02
CF300099	7LEAF--04
CF302576	7LEAF--08
CF301852	7LEAF--06
CF293115	30DGS--02
CF295077	30DGS--04
CF335603	JMT--05-F
CF294230	30DGS--03
CF299231	7LEAF--03
BM038487	V002D10 O
BF795494	H023H07 E
CF297684	30DGS--08
CF333485	JMT--02-G

C 98	16	72.7	437	7	CF300287	CF300287 7LEAF--04	C 171	16	72.7	491	7	CF335971	CF335971 JMT--05-N
C 99	16	72.7	438	7	CF336757	CF336757 JMT--06-O	C 172	16	72.7	492	7	CF335871	CF335871 JMT--05-L
C 100	16	72.7	439	7	CF336454	CF336454 JMT--06-I	C 173	16	72.7	493	7	CF302098	CF302098 7LEAF--07
C 101	16	72.7	440	4	BI810469	BI810469 J009B04 O	C 174	16	72.7	493	7	CF336032	CF336032 JMT--05-P
C 102	16	72.7	443	7	CF298634	CF298634 7LEAF--02	C 175	16	72.7	495	7	CF294456	CF294456 3ODGS--03
C 103	16	72.7	443	7	CF326605	CF326605 JMT1--06-F	C 176	16	72.7	495	7	CF301786	CF301786 7LEAF--06
C 104	16	72.7	444	7	CF296710	CF296710 3ODGS--07	C 177	16	72.7	495	7	CF336909	CF336909 JMT--07-C
C 105	16	72.7	447	7	CF300126	CF300126 7LEAF--04	C 178	16	72.7	496	7	CF292748	CF292748 3ODGS--01
C 106	16	72.7	447	7	CF334006	CF334006 JMT--03-C	C 179	16	72.7	496	7	CF299949	CF299949 7LEAF--04
C 107	16	72.7	447	7	CF335012	CF335012 JMT--04-I	C 180	16	72.7	496	7	CF336522	CF336522 JMT--06-J
C 108	16	72.7	448	7	CF292622	CF292622 3ODGS--01	C 181	16	72.7	497	7	CF300189	CF300189 7LEAF--04
C 109	16	72.7	448	7	CF298719	CF298719 7LEAF--02	C 182	16	72.7	497	7	CF335479	CF335479 JMT--05-C
C 110	16	72.7	452	7	CF299236	CF299236 7LEAF--02	C 183	16	72.7	498	7	CF294477	CF294477 3ODGS--04
C 111	16	72.7	454	7	CF307856	CF307856 ABF--01-H	C 184	16	72.7	498	7	CF297419	CF297419 3ODGS--08
C 112	16	72.7	455	7	CF298335	CF298335 7LEAF--01	C 185	16	72.7	499	5	BQ907486	BQ907486 P007C11 O
C 113	16	72.7	455	7	CF333435	CF333435 JMT--02-F	C 186	16	72.7	499	7	CF296926	CF296926 3ODGS--07
C 114	16	72.7	459	7	CF295831	CF295831 3ODGS--05	C 187	16	72.7	499	7	CF299465	CF299465 7LEAF--03
C 115	16	72.7	459	7	CF302341	CF302341 7LEAF--07	C 188	16	72.7	499	7	CF302481	CF302481 7LEAF--08
C 116	16	72.7	460	4	BI807599	BI807599 A006C09 O	C 189	16	72.7	499	7	CF332788	CF332788 JMT--01-G
C 117	16	72.7	461	7	CF195895	CF195895 A2b12.1 A	C 190	16	72.7	500	7	CF332942	CF332942 JMT--01-K
C 118	16	72.7	461	7	CF311539	CF311539 ABF--06-M	C 191	16	72.7	500	7	CF335956	CF335956 JMT--05-N
C 119	16	72.7	462	7	CF285701	CF285701 3ODGS--05	C 192	16	72.7	500	7	CF336502	CF336502 JMT--06-J
C 120	16	72.7	462	7	CF297362	CF297362 3ODGS--08	C 193	16	72.7	501	7	CF293284	CF293284 3ODGS--02
C 121	16	72.7	462	7	CF299260	CF299260 7LEAF--03	C 194	16	72.7	501	7	CF300694	CF300694 7LEAF--05
C 122	16	72.7	463	7	CF295676	CF295676 3ODGS--05	C 195	16	72.7	501	7	CF334025	CF334025 JMT--03-C
C 123	16	72.7	464	7	CF295075	CF295075 3ODGS--04	C 196	16	72.7	501	7	CF334152	CF334152 JMT--03-F
C 124	16	72.7	464	7	CF298102	CF298102 7LEAF--01	C 197	16	72.7	501	7	CF335291	CF335291 JMT--04-O
C 125	16	72.7	466	7	CF333098	CF333098 JMT--01-N	C 198	16	72.7	501	7	CF335896	CF335896 JMT--05-M
C 126	16	72.7	466	7	CF336355	CF336355 JMT--06-G	C 199	16	72.7	502	5	BQ529578	BQ529578 h12--03.8e
C 127	16	72.7	467	7	CF295382	CF295382 3ODGS--05	C 200	16	72.7	502	7	CF333698	CF333698 JMT--02-L
C 128	16	72.7	467	7	CF299470	CF299470 7LEAF--03	C 201	16	72.7	503	7	CF308990	CF308990 ABF--03-A
C 129	16	72.7	467	7	CF332819	CF332819 JMT--01-H	C 202	16	72.7	503	7	CF336371	CF336371 JMT--06-G
C 130	16	72.7	467	7	CF338041	CF338041 JMT--08-L	C 203	16	72.7	504	7	CF302018	CF302018 7LEAF--07
C 131	16	72.7	469	7	CF300208	CF300208 7LEAF--04	C 204	16	72.7	504	7	CF302471	CF302471 7LEAF--08
C 132	16	72.7	470	7	CF298227	CF298227 7LEAF--01	C 205	16	72.7	504	7	CF312560	CF312560 ABF--08-F
C 133	16	72.7	470	7	CF302203	CF302203 7LEAF--07	C 206	16	72.7	504	7	CF337598	CF337598 JMT--08-B
C 134	16	72.7	472	7	CF297460	CF297460 3ODGS--08	C 207	16	72.7	505	1	AA753360	AA753360 97AS2393
C 135	16	72.7	472	7	CF299893	CF299893 7LEAF--04	C 208	16	72.7	505	5	BQ907567	BQ907567 P008G11 O
C 136	16	72.7	472	7	CF333611	CF333611 JMT--02-J	C 209	16	72.7	505	7	CF301004	CF301004 7LEAF--05
C 137	16	72.7	473	7	CF334916	CF334916 JMT--04-G	C 210	16	72.7	505	7	CF335313	CF335313 JMT--04-P
C 138	16	72.7	474	7	CF311574	CF311574 ABF--06-N	C 211	16	72.7	505	7	CF337248	CF337248 JMT--07-J
C 139	16	72.7	474	7	CF334359	CF334359 JMT--03-K	C 212	16	72.7	506	7	CF296585	CF296585 3ODGS--07
C 140	16	72.7	474	7	CF335615	CF335615 JMT--05-F	C 213	16	72.7	506	7	CF333254	CF333254 JMT--02-B
C 141	16	72.7	477	4	BI809799	BI809799 G003A04 O	C 214	16	72.7	506	7	CF336913	CF336913 JMT--07-C
C 142	16	72.7	478	7	CF333330	CF333330 JMT--02-D	C 215	16	72.7	507	7	CF295579	CF295579 3ODGS--08
C 143	16	72.7	478	7	CF338043	CF338043 JMT--08-L	C 216	16	72.7	507	7	CF302633	CF302633 7LEAF--05
C 144	16	72.7	479	2	BF430561	BF430561 OGO2C02T3	C 217	16	72.7	507	7	CF334079	CF334079 JMT--03-D
C 145	16	72.7	479	7	CF302933	CF302933 7LEAF--08	C 218	16	72.7	508	7	CF333590	CF333590 JMT--02-I
C 146	16	72.7	480	4	BM419049	BM419049 R008F02 O	C 219	16	72.7	508	7	CF334927	CF334927 JMT--04-G
C 147	16	72.7	480	7	CF333127	CF333127 JMT--01-O	C 220	16	72.7	509	7	CF312061	CF312061 ABF--07-J
C 148	16	72.7	481	7	CF312849	CF312849 ABF--08-M	C 221	16	72.7	509	7	CF332812	CF332812 JMT--01-H
C 149	16	72.7	481	7	CF335083	CF335083 JMT--04-J	C 222	16	72.7	509	7	CF335309	CF335309 JMT--04-P
C 150	16	72.7	484	7	CF297887	CF297887 7LEAF--01	C 223	16	72.7	509	7	CF336057	CF336057 JMT--05-P
C 151	16	72.7	484	7	CF302376	CF302376 7LEAF--07	C 224	16	72.7	509	7	CF337639	CF337639 JMT--08-C
C 152	16	72.7	484	7	CF334479	CF334479 JMT--03-M	C 225	16	72.7	509	7	CF337734	CF337734 JMT--08-E
C 153	16	72.7	484	7	CF335350	CF335350 JMT--04-P	C 226	16	72.7	510	7	CF333207	CF333207 JMT--02-A
C 154	16	72.7	485	7	CF302641	CF302641 7LEAF--08	C 227	16	72.7	510	7	CF336446	CF336446 JMT--06-I
C 155	16	72.7	485	7	CF336484	CF336484 JMT--06-I	C 228	16	72.7	511	2	BE229988	BE229988 99AS225 R
C 156	16	72.7	486	7	CF295252	CF295252 3ODGS--05	C 229	16	72.7	511	7	CF300116	CF300116 7LEAF--04
C 157	16	72.7	486	7	CF297412	CF297412 3ODGS--08	C 230	16	72.7	511	7	CF337862	CF337862 JMT--08-H
C 158	16	72.7	486	7	CF299637	CF299637 7LEAF--03	C 231	16	72.7	512	7	CF294297	CF294297 3ODGS--03
C 159	16	72.7	486	7	CF336999	CF336999 JMT--07-E	C 232	16	72.7	512	7	CF297126	CF297126 3ODGS--07
C 160	16	72.7	486	7	CF337803	CF337803 JMT--08-O	C 233	16	72.7	513	7	CF299626	CF299626 7LEAF--03
C 161	16	72.7	487	7	CF302594	CF302594 7LEAF--08	C 234	16	72.7	513	7	CF336924	CF336924 JMT--07-C
C 162	16	72.7	487	7	CF332798	CF332798 JMT--01-H	C 235	16	72.7	514	7	CF295780	CF295780 3ODGS--05
C 163	16	72.7	487	7	CF334015	CF334015 JMT--03-C	C 236	16	72.7	514	7	CF298176	CF298176 7LEAF--01
C 164	16	72.7	487	7	CF335223	CF335223 JMT--04-N	C 237	16	72.7	514	7	CF302889	CF302889 7LEAF--08
C 165	16	72.7	487	7	CF335874	CF335874 JMT--05-L	C 238	16	72.7	514	7	CR286903	CR286903 CR286903
C 166	16	72.7	487	7	CF335893	CF335893 JMT--05-L	C 239	16	72.7	515	7	CF298884	CF298884 7LEAF--02
C 167	16	72.7	488	7	CF333761	CF333761 JMT--02-M	C 240	16	72.7	516	7	CF337361	CF337361 JMT--07-M
C 168	16	72.7	488	7	CF334521	CF334521 JMT--03-N	C 241	16	72.7	517	7	CF332599	CF332599 JMT--01-C
C 169	16	72.7	490	7	CF294832	CF294832 3ODGS--04	C 242	16	72.7	519	4	BM419793	BM419793 R018G12 O
C 170	16	72.7	490	7	CF335453	CF335453 JMT--05-C	C 243	16	72.7	520	4	BM419601	BM419601 R016B01 O

Tel: 82 331 290 0301
 Fax: 82 331 290 0307
 Email: myeun@sun20.asti.re.kr.
 Location/Qualifiers

FEATURES

source

1. .270
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Milyang 23"
 /db_xref="taxon:39946"
 /clone="98AS3025"
 /tissue_type="Immature Seed"
 /dev_stage="5 days after pollination"
 /lab_host="E. coli SOLR"
 /clone_lib="Rice Immature Seed Lambda ZAPII cDNA Library"
 /note="Vector: phuescript SK(+); Site 1: EcoRI; Site 2:
 XhoI; Directional cDNA library inserted into lambda ZAPII
 vector at 5' end with EcoRI and 3' end with Xho I site."

ORIGIN

Query Match 72.7%; Score 16; DB 2; Length 270;
 Best Local Similarity 75.0%; Pred. No. 37;
 Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CUCACUCGUGAGCUCT 21

Db 32 CTCACCTGAGCTCT 47

RESULT 6

LOCUS

CF295034/c 272 bp mRNA linear EST 14-AUG-2003
 DEFINITION 30DGS--04-M16.g1 Rice leaf plasmid cDNA library I (30DGS) Oryza
 sativa (japonica cultivar-group) cDNA clone 30DGS--04-M16, mRNA
 sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1. (bases 1 to 272)
 Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,
 Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.
 Large-scale Sequencing Analysis of Rice ESTs
 Unpublished (2003)
 Contact: Nahm B.H.
 Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
 of Bioscience and Bioinformatics, Myongji University
 Yongin, Kyeonggi, Korea
 Tel: 82 31 330 6193
 Fax: 82 31 321 6355
 Email: bhnam@gbio.com, bhnam@bio.myongji.ac.kr.

FEATURES

source

1. .272
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nackdong"
 /db_xref="taxon:39947"
 /clone="30DGS--04-M16"
 /tissue_type="leaf"
 /dev_stage="30 days after germination"
 /lab_host="E. coli DH10B"
 /clone_lib="Rice leaf plasmid cDNA library I (30DGS)"
 /note="vector: pCR4-TOP0; Site 1: EcoRI; mRNA was capped
 with oligoribonucleotides and then used as templates for
 RT-PCR."

ORIGIN

Query Match 72.7%; Score 16; DB 7; Length 272;
 Best Local Similarity 75.0%; Pred. No. 37;
 Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CUCACUCGUGAGCUCT 21

Db 129 CTCACCTGAGCTCT 114

RESULT 7

LOCUS

CF299062/c 273 bp mRNA linear EST 15-AUG-2003
 DEFINITION 7LEAF--02-O13.g1 Rice leaf plasmid cDNA library II (7LEAF) Oryza
 sativa (japonica cultivar-group) cDNA clone 7LEAF--02-O13, mRNA
 sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1. (bases 1 to 273)
 Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,
 Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.
 Large-scale Sequencing Analysis of Rice ESTs
 Unpublished (2003)
 Contact: Nahm B.H.
 Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
 of Bioscience and Bioinformatics, Myongji University
 Yongin, Kyeonggi, Korea
 Tel: 82 31 330 6193
 Fax: 82 31 321 6355
 Email: bhnam@gbio.com, bhnam@bio.myongji.ac.kr.

FEATURES

source

1. .273
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nackdong"
 /db_xref="taxon:39947"
 /clone="7LEAF--02-O13"
 /tissue_type="leaf"
 /dev_stage="7 days after germination"
 /lab_host="E. coli DH10B"
 /clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
 /note="vector: pCR4-TOP0; Site 1: EcoRI; mRNA was capped
 with oligoribonucleotides and then used as templates for
 RT-PCR."

ORIGIN

Query Match 72.7%; Score 16; DB 7; Length 273;
 Best Local Similarity 75.0%; Pred. No. 37;
 Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CUCACUCGUGAGCUCT 21

Db 182 CTCACCTGAGCTCT 167

RESULT 8

LOCUS

CF300491/c 273 bp mRNA linear EST 15-AUG-2003
 DEFINITION 7LEAF--04-P09.g1 Rice leaf plasmid cDNA library II (7LEAF) Oryza
 sativa (japonica cultivar-group) cDNA clone 7LEAF--04-P09, mRNA
 sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

AUTHORS

1. (bases 1 to 273)
 Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,


```

SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 (bases 1 to 281)
AUTHORS
Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,
Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.
TITLE
Large-scale Sequencing Analysis of Rice ESTs
JOURNAL
Unpublished (2003)
COMMENT
Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bnhahm@gbio.com, bnhahm@bio.myongji.ac.kr.
FEATURES
Location/Qualifiers
1..281
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39947"
/clone="7LEAF-01-P09"
/tissue_type="leaf"
/dev_stage="7 days after germination"
/lab_host="E.coli DH10B"
/notes="Vector: pCR4-TOPO; Site 1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."
ORIGIN
Query Match 72.7%; Score 16; DB 7; Length 281;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Qy 6 CUCACUCGUGAGCUCT 21
|:|:|:|:|:|:|
Db 115 CTCACCTCGTGAGCTCT 100

RESULT 12
CF294911/c
LOCUS
283 bp mRNA linear EST 14-AUG-2003
DEFINITION
30DGS--04-J22.g1 Rice leaf plasmid cDNA library I (30DGS) Oryza
sativa (japonica cultivar-group) cDNA clone 30DGS--04-J22, mRNA
sequence.
ACCESSION
CF294911 GI:33663944
VERSION
CF294911.1
KEYWORDS
Oryza sativa (japonica cultivar-group)
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 (bases 1 to 283)
AUTHORS
Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,
Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.
TITLE
Large-scale Sequencing Analysis of Rice ESTs
JOURNAL
Unpublished (2003)
COMMENT
Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bnhahm@gbio.com, bnhahm@bio.myongji.ac.kr.
FEATURES
Location/Qualifiers
1..283
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"

```

```

/db_xref="taxon:39947"
/clone="30DGS--04-J22"
/tissue_type="leaf"
/dev_stage="30 days after germination"
/lab_host="E.coli DH10B"
/clone_lib="Rice leaf plasmid cDNA library I (30DGS)"
/notes="Vector: pCR4-TOPO; Site 1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."
ORIGIN
Query Match 72.7%; Score 16; DB 7; Length 283;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Qy 6 CUCACUCGUGAGCUCT 21
|:|:|:|:|:|:|
Db 123 CTCACCTCGTGAGCTCT 108

RESULT 13
CF299992/c
LOCUS
311 bp mRNA linear EST 15-AUG-2003
DEFINITION
7LEAF--04-D16.g1 Rice leaf plasmid cDNA library II (7LEAF) Oryza
sativa (japonica cultivar-group) cDNA clone 7LEAF--04-D16, mRNA
sequence.
ACCESSION
CF299992 GI:33671753
VERSION
CF299992.1
KEYWORDS
Oryza sativa (japonica cultivar-group)
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 (bases 1 to 311)
AUTHORS
Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,
Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.
TITLE
Large-scale Sequencing Analysis of Rice ESTs
JOURNAL
Unpublished (2003)
COMMENT
Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bnhahm@gbio.com, bnhahm@bio.myongji.ac.kr.
FEATURES
Location/Qualifiers
1..311
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39947"
/clone="7LEAF--04-D16"
/tissue_type="leaf"
/dev_stage="7 days after germination"
/lab_host="E.coli DH10B"
/clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
/notes="Vector: pCR4-TOPO; Site 1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."
ORIGIN
Query Match 72.7%; Score 16; DB 7; Length 311;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Qy 6 CUCACUCGUGAGCUCT 21
|:|:|:|:|:|:|
Db 293 CTCACCTCGTGAGCTCT 278

RESULT 14
CB668085

```



```

Query Match      72.7%; Score 16; DB 7; Length 321;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy  6  CUCACUCGUGAGGCUCT 21
Db    261 CTCACCTCGTGAGCTCT 246

RESULT 17
CF334972/c
LOCUS
DEFINITION JMT--04-H11.g1 AtJMT-overexpressing transgenic rice plasmid cDNA
            library (JMT) Oryza sativa (japonica cultivar-group) cDNA clone
JMT--04-H11, mRNA sequence.
ACCESSION CF334972
VERSION   CF334972.1 GI:33818275
KEYWORDS EST.
SOURCE    Oryza sativa (japonica cultivar-group)
ORGANISM  Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE  1 (bases 1 to 321)
AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE     Large-scale Sequencing Analysis of Rice ESTs
JOURNAL   Unpublished (2003)
COMMENT   Contact: Nahm B.H.
           Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
           of Bioscience and Bioinformatics, Myongji University
           Yongin, Kyeonggi, Korea
           Tel: 82 31 330 6193
           Fax: 82 31 321 6355
           Email: bhna@mgbio.com, bhna@mgbio.myongji.ac.kr.

FEATURES             source
    Location/Qualifiers
        1..321
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="mRNA"
            /cultivar="Nackdong"
            /db_xref="taxon:39947"
            /clone="JMT--04-H11"
            /tissue_type="leaf"
            /dev_stage="14 days after germination"
            /lab_host="E.coli DH10B"
            /clone_lib="AtJMT-overexpressing transgenic rice plasmid
            cDNA library (JMT)"
            /note="vector: pCR4-TOPO; Site_1: EcoRI; Oligo-capped mRNA
            was reverse transcribed and then used for PCR. mRNA was
            prepared from Arabidopsis Jasmonate Carboxyl
            methyltransferase overexpression line."

ORIGIN
Query Match      72.7%; Score 16; DB 7; Length 321;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy  6  CUCACUCGUGAGGCUCT 21
Db    309 CTCACCTCGTGAGCTCT 294

RESULT 18
CF301775/c
LOCUS
DEFINITION 7LEAF--06-M05.g1 Rice leaf plasmid cDNA library II (7LEAF) Oryza
            sativa (japonica cultivar-group) cDNA clone 7LEAF--06-M05, mRNA
            sequence.
ACCESSION CF301775
VERSION   CF301775.1 GI:33673536
KEYWORDS EST.
SOURCE    Oryza sativa (japonica cultivar-group)

```

```

ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE  1 (bases 1 to 324)
AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE     Large-scale Sequencing Analysis of Rice ESTs
JOURNAL   Unpublished (2003)
COMMENT   Contact: Nahm B.H.
           Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
           of Bioscience and Bioinformatics, Myongji University
           Yongin, Kyeonggi, Korea
           Tel: 82 31 330 6193
           Fax: 82 31 321 6355
           Email: bhna@mgbio.com, bhna@mgbio.myongji.ac.kr.

FEATURES             source
    Location/Qualifiers
        1..324
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="mRNA"
            /cultivar="Nackdong"
            /db_xref="taxon:39947"
            /clone="7LEAF--06-M05"
            /tissue_type="leaf"
            /dev_stage="7 days after germination"
            /lab_host="E.coli DH10B"
            /clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
            /note="Vector: pCR4-TOPO; Site 1: EcoRI; mRNA was capped
            with oligoribonucleotides and then used as templates for
            RT-PCR."

ORIGIN
Query Match      72.7%; Score 16; DB 7; Length 324;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy  6  CUCACUCGUGAGGCUCT 21
Db    312 CTCACCTCGTGAGCTCT 297

RESULT 19
CF295092/c
LOCUS
DEFINITION CF295092 326 bp mRNA linear EST 14-AUG-2003
            30DGS--04-N22.b1 Rice leaf plasmid cDNA library I (30DGS) Oryza
            sativa (japonica cultivar-group) cDNA clone 30DGS--04-N22, mRNA
            sequence.
ACCESSION CF295092
VERSION   CF295092.1 GI:33664125
KEYWORDS EST.
SOURCE    Oryza sativa (japonica cultivar-group)
ORGANISM  Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE  1 (bases 1 to 326)
AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE     Large-scale Sequencing Analysis of Rice ESTs
JOURNAL   Unpublished (2003)
COMMENT   Contact: Nahm B.H.
           Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
           of Bioscience and Bioinformatics, Myongji University
           Yongin, Kyeonggi, Korea
           Tel: 82 31 330 6193
           Fax: 82 31 321 6355
           Email: bhna@mgbio.com, bhna@mgbio.myongji.ac.kr.

FEATURES             source
    Location/Qualifiers
        1..326
            /organism="Oryza sativa (japonica cultivar-group)"
            /mol_type="mRNA"
            /cultivar="Nackdong"
            /db_xref="taxon:39947"

```

```

/clone="30DGS--04-N22"
/tissue_type="leaf"
/dev_stage="30 days after germination"
/lab_host="E.coli DH10B"
/clone_lib="Rice leaf plasmid cDNA library I (30DGS)"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."

ORIGIN
Query Match          72.7%; Score 16; DB 7; Length 326;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      6 CUACUCUGAGGCUCT 21
        |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
Db      308 CTCACCTCGTGAGCTCT 293

RESULT 20
CF300479/c
LOCUS       CF300479.1   326 bp     mRNA           linear           EST 15-AUG-2003
DEFINITION   7LEAF--04-O17.g1 Rice leaf plasmid cDNA library II (7LEAF) Oryza sativa (japonica cultivar-group) cDNA clone 7LEAF--04-O17, mrna sequence.
ACCESSION    CF300479
VERSION      CF300479.1   GI:33672240
KEYWORDS     EST.
SOURCE       Oryza sativa (japonica cultivar-group)
ORGANISM     Oryza sativa (japonica cultivar-group)
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
              Ehrhartioideae; Oryzaceae; Oryza.
REFERENCE    1 (bases 1 to 326)
AUTHORS      Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
              Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE        Large-scale Sequencing Analysis of Rice ESTs
JOURNAL      Unpublished (2003)
COMMENT      Contact: Nahm B.H.
              Genomics and Bioinformatics, Myongji University
              of Bioscience and Biotechnology, GreenGene Biotech Inc.; Division
              Yongin, Kyeonggi, Korea
              Tel: 82 31 321 6355
              Fax: 82 31 321 6355
              Email: bnhahm@bio.com, bnhahm@bio.myongji.ac.kr.
              Location/Qualifiers
                  source
                      /organism="Oryza sativa (japonica cultivar-group)"
                      /mol_type="mrna"
                      /cultivar="Nackdong"
                      /db_xref="taxon:39947"
                      /clone="7LEAF--01-K01"
                      /tissue_type="leaf"
                      /dev_stage="7 days after germination"
                      /lab_host="E.coli DH10B"
                      /clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
                      /note="Vector: pCR4-TOPO; Site_1: EcoRI; mRNA was capped
                          with oligoribonucleotides and then used as templates for
                          RT-PCR."

ORIGIN
Query Match          72.7%; Score 16; DB 7; Length 329;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      6 CUACUCUGAGGCUCT 21
        |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
Db      311 CTCACCTCGTGAGCTCT 296

RESULT 22
CF301705/c
LOCUS       CF301705.1   330 bp     mRNA           linear           EST 15-AUG-2003
DEFINITION   7LEAF--06-K17.g1 Rice leaf plasmid cDNA library II (7LEAF) Oryza sativa (japonica cultivar-group) cDNA clone 7LEAF--06-K17, mrna sequence.
ACCESSION    CF301705
VERSION      CF301705.1   GI:33673466
KEYWORDS     EST.
SOURCE       Oryza sativa (japonica cultivar-group)
ORGANISM     Oryza sativa (japonica cultivar-group)
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
              Ehrhartioideae; Oryzaceae; Oryza.
REFERENCE    1 (bases 1 to 330)
AUTHORS      Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
              Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE        Large-scale Sequencing Analysis of Rice ESTs
JOURNAL      Unpublished (2003)
COMMENT      Contact: Nahm B.H.
              Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
              of Bioscience and Bioinformatics, Myongji University
              Yongin, Kyeonggi, Korea
              Tel: 82 31 330 6193
              Fax: 82 31 321 6355
              Email: bnhahm@bio.com, bnhahm@bio.myongji.ac.kr.
              Location/Qualifiers
                  source
                      /organism="Oryza sativa (japonica cultivar-group)"
                      /mol_type="mrna"
                      /cultivar="Nackdong"
                      /db_xref="taxon:39947"
                      /clone="7LEAF--04-O17"
                      /tissue_type="leaf"
                      /dev_stage="7 days after germination"
                      /lab_host="E.coli DH10B"
                      /clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
                      /note="Vector: pCR4-TOPO; Site_1: EcoRI; mRNA was capped
                          with oligoribonucleotides and then used as templates for
                          RT-PCR."

ORIGIN
Query Match          72.7%; Score 16; DB 7; Length 326;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      6 CUACUCUGAGGCUCT 21
        |:|:|:~
Db      308 CTCACCTCGTGAGCTCT 293

RESULT 21
CF298299/c
LOCUS       CF298299.1   329 bp     mRNA           linear           EST 15-AUG-2003
DEFINITION   7LEAF--01-K01.g1 Rice leaf plasmid cDNA library II (7LEAF) Oryza sativa (japonica cultivar-group) cDNA clone 7LEAF--01-K01, mrna sequence.
ACCESSION    CF298299
VERSION      CF298299.1   GI:33670060
KEYWORDS     EST.
SOURCE       Oryza sativa (japonica cultivar-group)
ORGANISM     Oryza sativa (japonica cultivar-group)
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
              Ehrhartioideae; Oryzaceae; Oryza.
REFERENCE    1 (bases 1 to 329)
AUTHORS      Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
              Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE        Large-scale Sequencing Analysis of Rice ESTs
JOURNAL      Unpublished (2003)
COMMENT      Contact: Nahm B.H.
              Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
              of Bioscience and Bioinformatics, Myongji University
              Yongin, Kyeonggi, Korea
              Tel: 82 31 330 6193
              Fax: 82 31 321 6355
              Email: bnhahm@bio.com, bnhahm@bio.myongji.ac.kr.
              Location/Qualifiers
                  source
                      /organism="Oryza sativa (japonica cultivar-group)"
                      /mol_type="mrna"
                      /cultivar="Nackdong"
                      /db_xref="taxon:39947"
                      /clone="7LEAF--01-K01"
                      /tissue_type="leaf"
                      /dev_stage="7 days after germination"
                      /lab_host="E.coli DH10B"
                      /clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
                      /note="Vector: pCR4-TOPO; Site_1: EcoRI; mRNA was capped
                          with oligoribonucleotides and then used as templates for
                          RT-PCR."

```


JOURNAL
COMMENT

Unpublished (2003)
Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bnhnm@gbio.com, bnhnm@bio.myongji.ac.kr.

FEATURES

source

1. .343
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39947"
/clone="7LEAF-05-H23"
/tissue_type="leaf"
/dev_stage="7 days after germination"
/lab_host="E.coli DH10B"
/clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
/notes="Vector: PCR4-TOPO; Site 1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."

ORIGIN

Query Match 72.7%; Score 16; DB 7; Length 343;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 6 CUCACUCGUGAGCUCT 21

|:|:|:|:|:|:|

Db 325 CTCACCTCGTGAGCTCT 310

RESULT 26

CF302308/c

LOCUS

DEFINITION
7LEAF--07-L05.b1 Rice leaf plasmid cDNA library II (7LEAF) Oryza
sativa (japonica cultivar-group) cDNA clone 7LEAF--07-L05, mRNA
sequence.

ACCESSION CF302308

VERSION 1 GI:33674069

KEYWORDS

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 349)

AUTHORS Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,
Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.

TITLE Large-scale Sequencing Analysis of Rice ESTs

JOURNAL Unpublished (2003)

COMMENT Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bnhnm@gbio.com, bnhnm@bio.myongji.ac.kr.

FEATURES

source

1. .349
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39947"
/clone="7LEAF-07-L05"
/tissue_type="leaf"
/dev_stage="7 days after germination"
/lab_host="E.coli DH10B"
/clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
/notes="Vector: PCR4-TOPO; Site 1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."

ORIGIN

Query Match 72.7%; Score 16; DB 7; Length 349;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 6 CUCACUCGUGAGCUCT 21

|:|:|:|:|:|:|

Db 331 CTCACCTCGTGAGCTCT 316

RESULT 27

BI810487

LOCUS

DEFINITION
BI810487 Oryza sativa mature leaf library induced by M.grisea Oryza
sativa cDNA clone J009E03, mRNA sequence.

ACCESSION BI810487

VERSION 1 GI:15857675

KEYWORDS

SOURCE Oryza sativa

ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 350)

AUTHORS Dong, H.T., Li, D.B., Zhuang, X.F., Dai, C.G., Sun, L.X., Pei, Y.X.,
Wu, H.F., Jiang, Y.X., Yu, P.C., Gao, Q.K. and Lou, Y.C.

TITLE A Gene Expression Screen in Oryza sativa

JOURNAL Unpublished (2001)

COMMENT Contact: Haitao Dong, Debao Li

Bioinformatics and Gene Network Research Group

Zhejiang University

Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China

Tel: 0086-571-86932051

Fax: 0086-571-86961525

Email: webmaster@estarray.org, URL: <http://www.estarray.org>

Seq primer: M3 forward primer.

FEATURES

source

1. .350
Location/Qualifiers
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="J009E03"
/tissue_type="leaf"
/dev_stage="Mature stage"
/clone_lib="Oryza sativa mature leaf library induced by
M.grisea"
/note="Vector: pSport2"

ORIGIN

Query Match 72.7%; Score 16; DB 4; Length 350;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 12; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 6 CUCACUCGUGAGCUCT 21

|:|:|:|:|:|:|

Db 11 CTCACCTCGTGAGCTCT 26

RESULT 28

CF333939/c

LOCUS

DEFINITION
JMT--03-A17.g1 AtJMT-overexpressing transgenic rice plasmid cDNA
library (JMT) Oryza sativa (japonica cultivar-group) cDNA clone

ACCESSION JMT--03-A17

VERSION 1 GI:33816183

KEYWORDS

SOURCE Oryza sativa

ORGANISM Oryza sativa

(japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

THIS PAGE IS BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 22:44:51 ; Search time 128 Seconds
(without alignments)
281.235 Million cell updates/sec

Title: US-10-848-737-1

Perfect score: 22

Sequence: 1 gugaacucacugagacuctt 22

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 300 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	15	68.2	234884	4	US-09-949-016-16420
2	14	63.6	348	3	Sequence 16420, A
3	14	63.6	428	4	Sequence 2127, Ap
4	14	63.6	428	4	Sequence 8994, Ap
5	14	63.6	3170	4	Sequence 24276, A
6	14	63.6	4329	4	Sequence 3849, Ap
7	14	63.6	300598	4	Sequence 2051, Ap
8	14	63.6	302604	4	Sequence 11868, A
9	14	63.6	302604	4	Sequence 14589, A
10	14	63.6	308362	4	Sequence 14589, A
11	14	63.6	4403765	3	Sequence 17119, A
12	14	63.6	4411529	3	Sequence 17119, A
13	13	59.1	240	4	US-09-103-840A-2
14	13	59.1	601	4	US-09-489-039A-216
15	13	59.1	601	4	Sequence 21714, A
16	13	59.1	601	4	Sequence 60781, A
17	13	59.1	601	4	Sequence 60782, A
18	13	59.1	601	4	Sequence 60783, A
19	13	59.1	601	4	Sequence 60784, A
20	13	59.1	1161	4	Sequence 164641, A
21	13	59.1	1411	4	Sequence 3861, Ap
22	13	59.1	1411	4	Sequence 505, App
23	13	59.1	1546	4	Sequence 4615, Ap
24	13	59.1	3250	3	Sequence 452, App
25	13	59.1	3935	3	Sequence 960, App
26	13	59.1	4082	4	Sequence 1, Appli
27	13	59.1	4082	4	Sequence 2722, Ap
					Sequence 50, Appli

US-09-949-016-12281	4	59.1	6831	13	28	Sequence 12281, A
US-09-949-016-14464	4	59.1	14205	13	29	Sequence 14464, A
US-09-949-016-12247	4	59.1	14395	13	30	Sequence 12247, A
US-09-949-016-16357	4	59.1	14395	13	31	Sequence 16357, A
US-09-949-016-17520	4	59.1	26709	13	32	Sequence 17520, A
US-09-341-587-6	3	59.1	29598	13	33	Sequence 6, Appli
US-09-949-016-17253	4	59.1	42672	13	34	Sequence 17253, A
US-09-949-016-17254	4	59.1	42672	13	35	Sequence 17254, A
US-09-335-409-1	3	59.1	68750	13	36	Sequence 1, Appli
US-09-568-102-1	3	59.1	68750	13	37	Sequence 1, Appli
US-09-567-969-1	3	59.1	68750	13	38	Sequence 1, Appli
US-09-568-480-1	3	59.1	68750	13	39	Sequence 1, Appli
US-09-568-486-1	3	59.1	68750	13	40	Sequence 1, Appli
US-09-568-479-1	3	59.1	68750	13	41	Sequence 1, Appli
US-09-567-899-1	3	59.1	68750	13	42	Sequence 1, Appli
US-09-949-016-15369	4	59.1	73757	13	43	Sequence 15369, A
US-09-949-016-14061	4	59.1	75216	13	44	Sequence 14061, A
US-09-949-016-16926	4	59.1	100928	13	45	Sequence 16926, A
US-09-949-016-12762	4	59.1	117937	13	46	Sequence 12762, A
US-09-949-016-15775	4	59.1	117937	13	47	Sequence 15775, A
US-09-949-016-13506	4	59.1	283538	13	48	Sequence 13506, A
US-09-557-884-1	4	59.1	1830121	13	49	Sequence 1, Appli
US-09-643-990A-1	4	59.1	1830121	13	50	Sequence 1, Appli
US-08-026-320A-10	23	54.5	23	12	51	Sequence 10, Appli
5496924-50	6	54.5	24	6	52	Patent No. 5496924
5496924-50	6	54.5	24	6	53	Patent No. 5496924
US-09-396-196G-72642	25	54.5	25	4	54	Sequence 72642, A
US-09-268-347-52	12	54.5	110	3	55	Sequence 52, Appli
US-09-513-999C-19813	12	54.5	212	4	56	Sequence 19813, A
US-09-513-999C-365	12	54.5	244	4	57	Sequence 365, App
US-09-489-039A-1763	285	54.5	285	4	58	Sequence 1763, Ap
US-09-313-294A-1964	12	54.5	288	4	59	Sequence 1964, Ap
US-09-513-999C-2476	340	54.5	340	4	60	Sequence 2476, Ap
US-09-621-976-18415	367	54.5	367	4	61	Sequence 18415, Ap
US-09-513-999C-2824	371	54.5	371	4	62	Sequence 2824, Ap
US-09-621-976-10841	446	54.5	446	4	63	Sequence 10841, A
US-09-621-976-18219	466	54.5	466	4	64	Sequence 18219, A
US-09-328-352-3457	510	54.5	510	4	65	Sequence 3457, Ap
US-09-615-192A-225	536	54.5	536	3	66	Sequence 225, App
US-08-998-416-122	545	54.5	545	3	67	Sequence 122, App
US-09-949-016-23423	601	54.5	601	4	68	Sequence 23423, A
US-09-949-016-29772	601	54.5	601	4	69	Sequence 29772, A
US-09-949-016-30889	601	54.5	601	4	70	Sequence 30889, A
US-09-949-016-33980	601	54.5	601	4	71	Sequence 33980, A
US-09-949-016-33981	601	54.5	601	4	72	Sequence 33981, A
US-09-949-016-39068	601	54.5	601	4	73	Sequence 39068, A
US-09-949-016-48740	601	54.5	601	4	74	Sequence 48740, A
US-09-949-016-66284	601	54.5	601	4	75	Sequence 66284, A
US-09-949-016-88078	601	54.5	601	4	76	Sequence 88078, A
US-09-949-016-88079	601	54.5	601	4	77	Sequence 88079, A
US-09-949-016-105259	601	54.5	601	4	78	Sequence 105259, A
US-09-949-016-121845	601	54.5	601	4	79	Sequence 121845, A
US-09-949-016-133334	601	54.5	601	4	80	Sequence 133334, A
US-09-949-016-134957	601	54.5	601	4	81	Sequence 134957, A
US-09-949-016-139340	601	54.5	601	4	82	Sequence 139340, A
US-09-949-016-143143	601	54.5	601	4	83	Sequence 143143, A
US-09-949-016-143144	601	54.5	601	4	84	Sequence 143144, A
US-09-949-016-145763	601	54.5	601	4	85	Sequence 145763, A
US-09-949-016-146031	601	54.5	601	4	86	Sequence 146031, A
US-09-949-016-146299	601	54.5	601	4	87	Sequence 146299, A
US-09-949-016-146879	601	54.5	601	4	88	Sequence 146879, A
US-09-949-016-162011	601	54.5	601	4	89	Sequence 162011, A
US-09-949-016-162013	601	54.5	601	4	90	Sequence 162013, A
US-09-949-016-162014	601	54.5	601	4	91	Sequence 162014, A
US-09-949-016-176608	601	54.5	601	4	92	Sequence 176608, A
US-09-949-016-182426	601	54.5	601	4	93	Sequence 182426, A
US-09-949-016-183700	601	54.5	601	4	94	Sequence 183700, A
US-09-949-016-183893	601	54.5	601	4	95	Sequence 183893, A
US-09-949-016-187991	601	54.5	601	4	96	Sequence 187991, A
US-09-949-016-199215	601	54.5	601	4	97	Sequence 199215, A
US-09-949-016-199216	601	54.5	601	4	98	Sequence 199216, A
US-08-998-416-119	606	54.5	606	3	99	Sequence 119, App
US-09-489-039A-5322	744	54.5	744	4	100	Sequence 5322, Ap

c 101	12	54.5	832	2	US-08-686-417-1	Sequence 1, Appl	174	12	54.5	3294	4	US-09-684-707-1	Sequence 1, Appl
102	12	54.5	1047	3	US-09-328-352-3329	Sequence 3329, Ap	175	12	54.5	3300	3	US-09-268-347-31	Sequence 31, Appl
103	12	54.5	1070	3	US-08-904-284-1	Sequence 1, Appl	c 176	12	54.5	3574	3	US-09-311-236-1	Sequence 1, Appl
c 104	12	54.5	1216	4	US-09-489-039A-6813	Sequence 6813, Ap	c 177	12	54.5	3574	4	US-09-912-917-1	Sequence 1, Appl
105	12	54.5	1251	4	US-09-328-352-1546	Sequence 1546, Ap	c 178	12	54.5	3609	3	US-09-705-299-11	Sequence 11, Appl
106	12	54.5	1260	4	US-09-602-777A-111	Sequence 111, App	c 179	12	54.5	3642	3	US-08-946-026-16	Sequence 16, Appl
c 107	12	54.5	1317	4	US-09-543-681A-2202	Sequence 2202, Ap	c 180	12	54.5	3720	3	US-09-342-681C-12	Sequence 12, Appl
c 108	12	54.5	1330	3	US-09-118-442-29	Sequence 29, Appl	c 181	12	54.5	3748	2	US-08-686-417-2	Sequence 2, Appl
c 109	12	54.5	1330	3	US-09-677-064-29	Sequence 29, Appl	c 182	12	54.5	3825	6	5310678-2	Patent No. 5310678
c 110	12	54.5	1439	2	US-08-495-695B-27	Sequence 27, Appl	c 183	12	54.5	3825	6	5310678-2	Patent No. 5310678
c 111	12	54.5	1439	5	PCT-US94-14436-27	Sequence 27, Appl	c 184	12	54.5	3857	4	US-09-319-497-14	Sequence 14, Appl
c 112	12	54.5	1534	2	US-08-588-983-1	Sequence 1, Appl	c 185	12	54.5	3879	3	US-09-638-715-35	Sequence 35, Appl
c 113	12	54.5	1534	2	US-08-588-976-1	Sequence 1, Appl	c 186	12	54.5	3879	3	US-10-060-509-35	Sequence 35, Appl
c 114	12	54.5	1645	2	US-08-724-394A-14	Sequence 14, Appl	c 187	12	54.5	3879	4	US-10-060-506-35	Sequence 35, Appl
c 115	12	54.5	1712	3	US-09-302-620B-105	Sequence 105, App	c 188	12	54.5	3948	4	US-09-799-451-125	Sequence 125, App
c 116	12	54.5	1712	4	US-09-912-161-45	Sequence 45, Appl	c 189	12	54.5	4177	2	US-08-484-575A-12	Sequence 12, Appl
c 117	12	54.5	1712	4	US-09-911-781-23	Sequence 23, Appl	c 190	12	54.5	4177	3	US-08-477-459-12	Sequence 12, Appl
c 118	12	54.5	1712	4	US-10-400-902-23	Sequence 23, Appl	c 191	12	54.5	4177	3	US-08-479-869-12	Sequence 12, Appl
c 119	12	54.5	1722	4	US-09-949-016-3155	Sequence 3155, Ap	c 192	12	54.5	4177	3	US-08-486-414-12	Sequence 12, Appl
c 120	12	54.5	1722	4	US-09-949-016-3241	Sequence 3241, Ap	c 193	12	54.5	4177	5	PCT-US94-01826A-12	Sequence 12, Appl
c 121	12	54.5	1734	2	US-08-663-566A-10	Sequence 10, Appl	c 194	12	54.5	4177	5	PCT-US94-02252A-12	Sequence 12, Appl
c 122	12	54.5	1734	2	US-08-023-610-10	Sequence 10, Appl	c 195	12	54.5	4463	2	US-08-760-489-1	Sequence 1, Appl
c 123	12	54.5	1734	2	US-08-288-065A-10	Sequence 10, Appl	c 196	12	54.5	4463	2	US-08-760-489-3	Sequence 3, Appl
c 124	12	54.5	1734	2	US-08-362-240A-10	Sequence 10, Appl	c 197	12	54.5	4463	3	US-09-185-373-1	Sequence 1, Appl
c 125	12	54.5	1734	3	US-08-804-372A-8	Sequence 8, Appl	c 198	12	54.5	4463	3	US-09-185-373-3	Sequence 3, Appl
c 126	12	54.5	1734	5	PCT-US95-10245-10	Sequence 10, Appl	c 199	12	54.5	4902	4	US-09-902-540-2065	Sequence 2065, Ap
c 127	12	54.5	1759	4	US-09-667-135-3	Sequence 3, Appl	c 200	12	54.5	5524	4	US-09-844-497-3	Sequence 3, Appl
c 128	12	54.5	1871	4	US-09-949-016-3156	Sequence 3156, Ap	c 201	12	54.5	5738	1	US-08-409-995-3	Sequence 3, Appl
c 129	12	54.5	1871	4	US-09-949-016-3242	Sequence 3242, Ap	c 202	12	54.5	5738	3	US-08-685-467-3	Sequence 3, Appl
c 130	12	54.5	1907	1	US-07-820-154A-29	Sequence 29, Appl	c 203	12	54.5	6129	4	US-09-489-039A-6882	Sequence 6882, Ap
c 131	12	54.5	1907	3	US-08-097-554A-29	Sequence 29, Appl	c 204	12	54.5	6678	3	US-09-268-347-33	Sequence 33, Appl
c 132	12	54.5	1907	3	US-08-480-640A-29	Sequence 29, Appl	c 205	12	54.5	6706	3	US-09-268-347-27	Sequence 27, Appl
c 133	12	54.5	1907	3	US-08-295-802-29	Sequence 29, Appl	c 206	12	54.5	6766	4	US-09-527-058-1	Sequence 1, Appl
c 134	12	54.5	1907	3	US-08-488-237A-29	Sequence 29, Appl	c 207	12	54.5	7152	4	US-09-492-709A-142	Sequence 142, App
c 135	12	54.5	1907	3	US-08-375-392A-29	Sequence 29, Appl	c 208	12	54.5	7253	3	US-09-268-347-35	Sequence 35, Appl
c 136	12	54.5	1907	4	US-08-472-679B-29	Sequence 29, Appl	c 209	12	54.5	7291	3	US-08-513-942-3	Sequence 3, Appl
c 137	12	54.5	1907	5	PCT-US93-00324-29	Sequence 29, Appl	c 210	12	54.5	7291	3	US-09-684-707-3	Sequence 3, Appl
c 138	12	54.5	2037	3	US-08-913-942-14	Sequence 14, Appl	c 211	12	54.5	7432	4	US-09-949-016-16414	Sequence 16414, A
c 139	12	54.5	2037	4	US-09-684-707-14	Sequence 14, Appl	c 212	12	54.5	9064	3	US-08-961-527-17	Sequence 17, Appl
c 140	12	54.5	2079	3	US-09-268-347-25	Sequence 25, Appl	c 213	12	54.5	10288	4	US-09-949-016-14897	Sequence 14897, A
c 141	12	54.5	2154	4	US-09-583-110-1968	Sequence 1968, Ap	c 214	12	54.5	10288	4	US-09-949-016-14893	Sequence 14893, A
c 142	12	54.5	2166	4	US-09-107-433-216	Sequence 216, Ap	c 215	12	54.5	10303	4	US-09-949-016-14898	Sequence 14898, A
c 143	12	54.5	2307	4	US-09-786-480B-1	Sequence 1, Appl	c 216	12	54.5	10304	4	US-09-949-016-14984	Sequence 14984, A
c 144	12	54.5	2417	3	US-09-439-313-334	Sequence 334, App	c 217	12	54.5	11766	4	US-09-949-016-12531	Sequence 12531, A
c 145	12	54.5	2417	3	US-09-352-616A-334	Sequence 334, App	c 218	12	54.5	11770	4	US-09-949-016-12720	Sequence 12720, A
c 146	12	54.5	2417	3	US-09-332-149A-334	Sequence 334, App	c 219	12	54.5	11770	4	US-09-949-016-13487	Sequence 13487, A
c 147	12	54.5	2417	4	US-09-636-215-334	Sequence 334, App	c 220	12	54.5	11770	4	US-09-949-016-13488	Sequence 13488, A
c 148	12	54.5	2417	4	US-09-685-166A-334	Sequence 334, App	c 221	12	54.5	12249	4	US-09-949-016-16839	Sequence 16839, A
c 149	12	54.5	2417	4	US-09-688-489-334	Sequence 334, App	c 222	12	54.5	12249	4	US-09-949-016-16840	Sequence 16840, A
c 150	12	54.5	2417	4	US-09-679-426-334	Sequence 334, App	c 223	12	54.5	13795	4	US-09-949-016-15323	Sequence 15323, A
c 151	12	54.5	2417	4	US-09-759-143-334	Sequence 334, App	c 224	12	54.5	14519	4	US-09-949-016-15611	Sequence 15611, A
c 152	12	54.5	2417	4	US-09-651-236-334	Sequence 334, App	c 225	12	54.5	14519	4	US-09-949-016-15612	Sequence 15612, A
c 153	12	54.5	2433	4	US-09-300-958A-24	Sequence 24, Appl	c 226	12	54.5	15186	4	US-09-741-744A-134	Sequence 134, App
c 154	12	54.5	2437	3	US-08-904-284-6	Sequence 6, Appl	c 227	12	54.5	15666	4	US-09-949-016-15929	Sequence 15929, A
c 155	12	54.5	2481	1	US-08-290-301-1	Sequence 1, Appl	c 228	12	54.5	16572	4	US-09-949-016-12409	Sequence 12409, A
c 156	12	54.5	2481	3	US-09-013-598-1	Sequence 1, Appl	c 229	12	54.5	16578	4	US-09-949-016-14680	Sequence 14680, A
c 157	12	54.5	2484	1	US-07-841-997A-1	Sequence 1, Appl	c 230	12	54.5	17353	4	US-09-949-016-13437	Sequence 13437, A
c 158	12	54.5	2718	3	US-09-651-656-14	Sequence 14, Appl	c 231	12	54.5	18159	4	US-09-949-016-12401	Sequence 12401, A
c 159	12	54.5	2718	3	US-09-650-855-14	Sequence 14, Appl	c 232	12	54.5	18160	4	US-09-949-016-13677	Sequence 13677, A
c 160	12	54.5	2860	3	US-09-221-017B-871	Sequence 871, App	c 233	12	54.5	19523	4	US-09-949-016-16344	Sequence 16344, A
c 161	12	54.5	2861	3	US-09-398-395A-55	Sequence 55, Appl	c 234	12	54.5	19826	4	US-09-949-016-16973	Sequence 16973, A
c 162	12	54.5	2861	4	US-09-887-586A-55	Sequence 55, Appl	c 235	12	54.5	20373	4	US-09-949-016-12302	Sequence 12302, A
c 163	12	54.5	2861	4	US-09-895-752-55	Sequence 55, Appl	c 236	12	54.5	20374	4	US-09-949-016-17086	Sequence 17086, A
c 164	12	54.5	2861	4	US-09-903-012B-55	Sequence 55, Appl	c 237	12	54.5	22080	4	US-09-949-016-16183	Sequence 16183, A
c 165	12	54.5	2861	4	US-09-900-797-55	Sequence 55, Appl	c 238	12	54.5	22807	4	US-09-902-540-1214	Sequence 1214, Ap
c 166	12	54.5	2968	4	US-09-527-058-6	Sequence 6, Appl	c 239	12	54.5	24663	4	US-09-949-016-14668	Sequence 14668, A
c 167	12	54.5	3030	3	US-09-268-347-29	Sequence 29, Appl	c 240	12	54.5	26454	4	US-09-902-540-1251	Sequence 1251, Ap
c 168	12	54.5	3036	3	US-09-268-347-23	Sequence 23, Appl	c 241	12	54.5	26709	4	US-09-949-016-17520	Sequence 17520, A
c 169	12	54.5	3115	3	US-09-231-017B-849	Sequence 849, App	c 242	12	54.5	28198	4	US-09-949-016-12349	Sequence 12349, A
c 170	12	54.5	3294	1	US-08-409-995-1	Sequence 1, Appl	c 243	12	54.5	35060	3	US-08-814-095-7	Sequence 7, Appl
c 171	12	54.5	3294	3	US-08-685-467-1	Sequence 1, Appl	c 244	12	54.5	35100	1	US-08-306-691B-19	Sequence 19, Appl
c 172	12	54.5	3294	3	US-08-913-942-1	Sequence 1, Appl	c 245	12	54.5	35100	5	PCT-US93-06251-19	Sequence 19, Appl
c 173	12	54.5	3294	3	US-09-268-347-43	Sequence 43, Appl	c 246	12	54.5	35707	4	US-09-949-016-17120	Sequence 17120, A

```

c 247 12 54.5 37059 4 US-09-949-016-15865 Sequence 15865, A-
c 248 12 54.5 40181 4 US-09-949-016-17016 Sequence 17016, A
c 249 12 54.5 40575 4 US-09-949-016-17246 Sequence 17246, A
c 250 12 54.5 42975 4 US-09-949-016-11965 Sequence 11965, A
c 251 12 54.5 43804 3 US-09-171-461-1 Sequence 1, Appli
c 252 12 54.5 43804 4 US-09-970-711-1 Sequence 1, Appli
c 253 12 54.5 44998 4 US-09-949-016-12824 Sequence 12824, A
c 254 12 54.5 44999 4 US-09-949-016-14832 Sequence 14832, A
c 255 12 54.5 44999 4 US-09-949-016-14833 Sequence 14833, A
c 256 12 54.5 44999 4 US-09-949-016-17108 Sequence 17108, A
c 257 12 54.5 44999 4 US-09-949-016-17109 Sequence 17109, A
c 258 12 54.5 45587 4 US-09-949-016-15836 Sequence 15836, A
c 259 12 54.5 47199 4 US-09-949-016-12144 Sequence 12144, A
c 260 12 54.5 47200 4 US-09-949-016-13526 Sequence 13526, A
c 261 12 54.5 51552 4 US-09-733-294A-30 Sequence 30, Appli
c 262 12 54.5 51928 4 US-09-949-016-13184 Sequence 13184, A
c 263 12 54.5 55703 4 US-09-949-016-12007 Sequence 12007, A
c 264 12 54.5 55703 4 US-09-949-016-16781 Sequence 16781, A
c 265 12 54.5 55841 4 US-09-949-016-16602 Sequence 16602, A
c 266 12 54.5 57218 4 US-09-949-016-16512 Sequence 16512, A
c 267 12 54.5 57331 4 US-09-949-016-17277 Sequence 17277, A
c 268 12 54.5 58782 4 US-09-949-016-16851 Sequence 16851, A
c 269 12 54.5 61399 4 US-09-949-016-14386 Sequence 14386, A
c 270 12 54.5 63804 4 US-09-949-016-15200 Sequence 15200, A
c 271 12 54.5 64046 4 US-09-949-016-12560 Sequence 12560, A
c 272 12 54.5 64047 4 US-09-949-016-15781 Sequence 15781, A
c 273 12 54.5 64610 4 US-09-949-016-12214 Sequence 12214, A
c 274 12 54.5 72128 4 US-09-949-016-16018 Sequence 16018, A
c 275 12 54.5 73295 4 US-09-949-016-15151 Sequence 15151, A
c 276 12 54.5 77772 4 US-09-949-016-17417 Sequence 17417, A
c 277 12 54.5 77997 4 US-09-949-016-12249 Sequence 12249, A
c 278 12 54.5 80269 4 US-09-949-016-15681 Sequence 15681, A
c 279 12 54.5 85369 4 US-09-949-016-12171 Sequence 12171, A
c 280 12 54.5 87780 4 US-09-949-016-17011 Sequence 17011, A
c 281 12 54.5 90428 4 US-09-949-016-12564 Sequence 12564, A
c 282 12 54.5 91831 4 US-09-949-016-13694 Sequence 13694, A
c 283 12 54.5 113100 4 US-09-949-016-12245 Sequence 12245, A
c 284 12 54.5 113701 4 US-09-949-016-13214 Sequence 13214, A
c 285 12 54.5 113876 4 US-09-949-016-14828 Sequence 14828, A
c 286 12 54.5 113876 4 US-09-949-016-14829 Sequence 14829, A
c 287 12 54.5 115508 4 US-09-949-016-11800 Sequence 11800, A
c 288 12 54.5 115508 4 US-09-949-016-14826 Sequence 14826, A
c 289 12 54.5 115508 4 US-09-949-016-14827 Sequence 14827, A
c 290 12 54.5 134434 4 US-09-949-016-17362 Sequence 17362, A
c 291 12 54.5 141115 4 US-09-949-016-17490 Sequence 17490, A
c 292 12 54.5 142504 4 US-09-949-016-13693 Sequence 13693, A
c 293 12 54.5 142506 4 US-09-949-016-12474 Sequence 12474, A
c 294 12 54.5 152070 4 US-09-949-016-15402 Sequence 15402, A
c 295 12 54.5 153866 4 US-09-949-016-16919 Sequence 16919, A
c 296 12 54.5 165651 4 US-09-949-016-13032 Sequence 13032, A
c 297 12 54.5 168971 4 US-09-949-016-13807 Sequence 13807, A
c 298 12 54.5 183202 4 US-09-949-016-13614 Sequence 13614, A
c 299 12 54.5 183770 4 US-09-949-016-15494 Sequence 15494, A
c 300 12 54.5 183770 4 US-09-949-016-15494 Sequence 15494, A

ALIGNMENTS

RESULT 1
US-09-949-016-16420/c
; Sequence 16420, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

Sequence 15865, A-
Sequence 17016, A
Sequence 17246, A
Sequence 11965, A
Sequence 1, Appli
Sequence 1, Appli
Sequence 12824, A
Sequence 14832, A
Sequence 14833, A
Sequence 17108, A
Sequence 17109, A
Sequence 15836, A
Sequence 12144, A
Sequence 13526, A
Sequence 30, Appli
Sequence 13184, A
Sequence 12007, A
Sequence 16781, A
Sequence 16602, A
Sequence 16512, A
Sequence 17277, A
Sequence 16851, A
Sequence 14386, A
Sequence 15200, A
Sequence 12560, A
Sequence 15781, A
Sequence 12214, A
Sequence 16018, A
Sequence 15151, A
Sequence 17417, A
Sequence 12249, A
Sequence 15681, A
Sequence 12171, A
Sequence 17011, A
Sequence 12564, A
Sequence 13694, A
Sequence 12245, A
Sequence 13214, A
Sequence 14828, A
Sequence 14829, A
Sequence 11800, A
Sequence 14826, A
Sequence 14827, A
Sequence 17362, A
Sequence 17490, A
Sequence 13693, A
Sequence 12474, A
Sequence 15402, A
Sequence 16919, A
Sequence 13032, A
Sequence 13807, A
Sequence 13614, A
Sequence 15494, A
Sequence 15494, A

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16420
; LENGTH: 234884
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(234884)
; OTHER INFORMATION: n = A, T, C or G
US-09-949-016-16420

Query Match 68.2%; Score 15; DB 4; Length 234884;
Best Local Similarity 80.0%; Pred. No. 5.2;
Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 8 CACUCUGAGCUCTT 22
||||:||||:||||
Db 140861 CACTCGAGCTCTT 140847

RESULT 2
US-09-134-001C-2127
; Sequence 2127, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2127
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2127

Query Match 63.6%; Score 14; DB 3; Length 348;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUGAACUCACUCGU 14
||||:||||:||||
Db 134 GTGAACCTCACTCT 147

RESULT 3
US-09-270-767-8994
; Sequence 8994, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8994
; LENGTH: 428
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-8994
```

Query Match 63.6%; Score 14; DB 4; Length 428;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 UCACUCGUGAGCUC 20
Db 356 TCACGCTGAGCTC 369
:||||:|||||

RESULT 4

US-09-270-767-24276
; Sequence 24276, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Honburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24276
; LENGTH: 428
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-24276

Query Match 63.6%; Score 14; DB 4; Length 428;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 UCACUCGUGAGCUC 20
Db 356 TCACGCTGAGCTC 369
:||||:|||||

RESULT 5

US-09-710-279-3849
; Sequence 3849, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3849
; LENGTH: 3170
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3849

Query Match 63.6%; Score 14; DB 4; Length 3170;
Best Local Similarity 71.4%; Pred. No. 24;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUGAACUCACUCGU 14
Db 1090 GTGAACCTCACTCGT 1103
|:||||:||||:

RESULT 6

US-09-710-279-2051
; Sequence 2051, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:

; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2051
; LENGTH: 4329
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-2051

Query Match 63.6%; Score 14; DB 4; Length 4329;
Best Local Similarity 71.4%; Pred. No. 24;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUGAACUCACUCGU 14
Db 101 GTGAACCTCACTCGT 114
|:||||:||||:

RESULT 7

US-09-949-016-11868/c
; Sequence 11868, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11868
; LENGTH: 300598
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(300598)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11868

Query Match 63.6%; Score 14; DB 4; Length 300598;
Best Local Similarity 78.6%; Pred. No. 21;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 9 ACUCGUGAGCUCTT 22
Db 93184 ACTCGTGAGCTCTT 93171
|:||||:||||:

RESULT 8

US-09-949-016-14588/c
; Sequence 14588, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14588
; LENGTH: 302604
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(302604)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14588

Query Match 63.6%; Score 14; DB 4; Length 302604;
Best Local Similarity 78.6%; Pred. No. 21;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 9 ACUCGUGAGCUCTT 22
Db 143190 ACTCGTGAGCTCTT 143177
||:|||||:

RESULT 9
US-09-949-016-14589/c
; Sequence 14589, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14589
; LENGTH: 302604
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(302604)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14589

Query Match 63.6%; Score 14; DB 4; Length 302604;
Best Local Similarity 78.6%; Pred. No. 21;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 9 ACUCGUGAGCUCTT 22
Db 143190 ACTCGTGAGCTCTT 143177
||:|||||:

RESULT 10
US-09-949-016-17119/c
; Sequence 17119, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17119
; LENGTH: 308362
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(308362)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17119

Query Match 63.6%; Score 14; DB 4; Length 308362;
Best Local Similarity 78.6%; Pred. No. 21;
Matches 11; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 9 ACUCGUGAGCUCTT 22
Db 143006 ACTCGTGAGCTCTT 142993
||:|||||:

RESULT 11
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: PRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 63.6%; Score 14; DB 3; Length 4403765;
Best Local Similarity 71.4%; Pred. No. 18;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 6 CUCACUCGUGAGCU 19
Db 786568 CTCACGAGGCT 786555
||:|||||:

RESULT 12
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:

<p>; APPLICANT: FLEISCHMAN, Robert D. ; APPLICANT: WHITE, Owen R. ; APPLICANT: FRASER, Claire M. ; APPLICANT: VENTER, John C. ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM ; TITLE OF INVENTION: TUBERCULOSIS ; FILE REFERENCE: 24366-2007.00 ; CURRENT APPLICATION NUMBER: US/09/103,840A ; CURRENT FILING DATE: 1998-06-24 ; NUMBER OF SEQ ID NOS: 2 ; SOFTWARE: PatentIn Ver. 2.1 ; SEQ ID NO 1 ; LENGTH: 4411529 ; TYPE: DNA ; ORGANISM: Mycobacterium tuberculosis ; OTHER INFORMATION: H37Rv US-09-103-840A-1</p>	<p>Query Match 63.6%; Score 14; DB 3; Length 4411529; Best Local Similarity 71.4%; Pred. No. 18; Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;</p>
<p>Qy 6 CUCACUCGUGAGCU 19 : :: :: : Db 784607 CTCCTCGTGAGCT 784594</p>	
<p>RESULT 13 US-09-489-039A-216 ; Sequence 216, Application US/09489039A ; Patent No. 6610836 ; GENERAL INFORMATION: ; APPLICANT: Gary Breton et al. ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS ; FILE REFERENCE: 2709.2004001 ; CURRENT APPLICATION NUMBER: US/09/489,039A ; CURRENT FILING DATE: 2000-01-27 ; PRIOR APPLICATION NUMBER: US 60/117,747 ; PRIOR FILING DATE: 1999-01-29 ; NUMBER OF SEQ ID NOS: 14342 ; SEQ ID NO 216 ; LENGTH: 240 ; TYPE: DNA ; ORGANISM: Klebsiella pneumoniae US-09-489-039A-216</p>	
<p>Query Match 59.1%; Score 13; DB 4; Length 240; Best Local Similarity 76.9%; Pred. No. 1.le+02; Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;</p>	
<p>Qy 6 CUCACUCGUGAGC 18 : :: :: : Db 129 CTCCTCGTGAGC 141</p>	
<p>RESULT 14 US-09-949-016-27714 ; Sequence 27714, Application US/09949016 ; Patent No. 6812339 ; GENERAL INFORMATION: ; APPLICANT: VENTER, J. Craig et al. ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF ; FILE REFERENCE: CL001307 ; CURRENT APPLICATION NUMBER: US/09/949,016 ; CURRENT FILING DATE: 2000-04-14 ; PRIOR APPLICATION NUMBER: 60/241,755 ; PRIOR FILING DATE: 2000-10-20 ; PRIOR APPLICATION NUMBER: 60/237,768 ; PRIOR FILING DATE: 2000-10-03 ; PRIOR APPLICATION NUMBER: 60/231,498 ; PRIOR FILING DATE: 2000-09-08 ; NUMBER OF SEQ ID NOS: 207012 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 60782 ; LENGTH: 601 ; TYPE: DNA</p>	

```

; ORGANISM: Human
US-09-949-016-60782

Query Match          59.1%; Score 13; DB 4; Length 601;
Best Local Similarity 76.9%; Pred. No. 1e+02;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCG 13
Db 339 GTGAACCTCACTCG 327

RESULT 17
US-09-949-016-60783/c
; Sequence 60783, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60783
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-60783

Query Match          59.1%; Score 13; DB 4; Length 601;
Best Local Similarity 76.9%; Pred. No. 1e+02;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCG 13
Db 332 GTGAACCTCACTCG 320

RESULT 18
US-09-949-016-60784/c
; Sequence 60784, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60784
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-60784

Query Match          59.1%; Score 13; DB 4; Length 601;
Best Local Similarity 76.9%; Pred. No. 1e+02;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCG 13
Db 233 GTGAACCTCACTCG 221

RESULT 19
US-09-949-016-164641
; Sequence 164641, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164641
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-164641

Query Match          59.1%; Score 13; DB 4; Length 601;
Best Local Similarity 76.9%; Pred. No. 1e+02;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AACUCACUCGUGA 16
Db 238 AACCTCACTCGTGA 250

RESULT 20
US-09-543-681A-3861
; Sequence 3861, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3861
; LENGTH: 1161
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-3861

Query Match          59.1%; Score 13; DB 4; Length 1161;
Best Local Similarity 69.2%; Pred. No. 1e+02;
Matches 9; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 7 UCACUCGUGAGCU 19
Db 170 TCACCTCGTGAGCT 182

RESULT 21
US-09-949-016-505
; Sequence 505, Application US/09949016

```

```
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 505
; LENGTH: 1411
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-505

Query Match          59.1%; Score 13; DB 4; Length 1411;
Best Local Similarity 76.9%; Pred. No. 1e+02;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      4 AACUCACUCGUGA 16
      |||:||||:||||
Db      453 AACTCACTCGTGA 465

RESULT 22
US-09-949-016-4615
; Sequence 4615, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4615
; LENGTH: 1411
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4615

Query Match          59.1%; Score 13; DB 4; Length 1411;
Best Local Similarity 76.9%; Pred. No. 1e+02;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      4 AACUCACUCGUGA 16
      |||:||||:||||
Db      453 AACTCACTCGTGA 465

RESULT 23
US-09-620-312D-452
; Sequence 452, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
```

```
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghaast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 452
; LENGTH: 1546
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (404)..(1189)
US-09-620-312D-452

Query Match          59.1%; Score 13; DB 4; Length 1546;
Best Local Similarity 76.9%; Pred. No. 99;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      4 AACUCACUCGUGA 16
      |||:||||:||||
Db      572 AACTCACTCGTGA 584

RESULT 24
US-09-221-017B-960/c
; Sequence 960, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PFI182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA: PFI546
; APPLICATION NUMBER: PFI546
; FILING DATE: 30-JAN-1998
```

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PP2911
;; FILING DATE: 09-APR-1998
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/AU98/01023
;; FILING DATE: 10-DEC-1998
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Monroy, Gladys H
;; REGISTRATION NUMBER: 32,430
;; REFERENCE/DOCKET NUMBER: 27340-20021.00
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650-813-5600
;; TELEFAX: 650-494-0792
;; TELEX: 706141
;; INFORMATION FOR SEQ ID NO: 960:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3250 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: circular
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ANTI-SENSE: UNKNOWN
;; ORIGINAL SOURCE:
;; ORGANISM: PORPHYROMONAS GINGIVALIS
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: 1...3250
US-09-221-017B-960

Query Match 59.1%; Score 13; DB 3; Length 3250;
Best Local Similarity 76.9%; Pred. No. 97;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCG 13
|:||||:||||:
Db 2826 GTGAACCTCACTCG 2814

RESULT 25
US-09-060-482-1/c
; Sequence 1, Application US/09060482
; Patent No. 6468766
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Layne, Matthew D.
; APPLICANT: Yet, Shaw-Fang
; TITLE OF INVENTION: AORTIC CARDIOPOLYPEPTIDE-LIKE POLYPEPTIDE
; FILE REFERENCE: 03433/036001
; CURRENT APPLICATION NUMBER: US/09/060,482
; CURRENT FILING DATE: 1998-04-15
; EARLIER APPLICATION NUMBER: US 08/818,009
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: US 60/013,439
; EARLIER FILING DATE: 1996-03-15
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3935
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (140)...(3613)
US-09-060-482-1

Query Match 59.1%; Score 13; DB 3; Length 3935;
Best Local Similarity 76.9%; Pred. No. 96;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 10 CUCGUGAGCUCTT 22
|:||||:||||:
Db 1925 CTCGTGAGCTCTT 1913

RESULT 26

US-09-949-016-2722/c
; Sequence 2722, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 2722
; LENGTH: 4082
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2722

Query Match 59.1%; Score 13; DB 4; Length 4082;
Best Local Similarity 76.9%; Pred. No. 96;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 10 CUCGUGAGCUCTT 22
|:||||:||||:
Db 2092 CTCGTGAGCTCTT 2080

RESULT 27

US-09-525-305-50
; Sequence 50, Application US/09525305
; Patent No. 6806062
; GENERAL INFORMATION:
; APPLICANT: Hjort, Carsten
; APPLICANT: Hondel, C.A.M.J.J. van den
; APPLICANT: Punt, P.J.
; APPLICANT: Shuren, F.H.J.
; APPLICANT: Christensen, Tove
; TITLE OF INVENTION: Fungal Transcriptional Activator Useful In Methods For Producing
; Polypeptides
; FILE REFERENCE: 5555.400-US
; CURRENT APPLICATION NUMBER: US/09/525,305
; CURRENT FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 50
; LENGTH: 6015
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pdv8 plasmid
US-09-525-305-50

Query Match 59.1%; Score 13; DB 4; Length 6015;
Best Local Similarity 76.9%; Pred. No. 95;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 9 ACUCGUGAGCUCT 21
|:||||:||||:
Db 350 ACTCGTGAAGCTCT 362

RESULT 28

US-09-949-016-12281
; Sequence 12281, Application US/09949016

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12281
; LENGTH: 6831
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12281

Query Match 59.1%; Score 13; DB 4; Length 6831;
Best Local Similarity 69.2%; Pred. No. 95;
Matches 9; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 7 UCACUGUGAGCU 19
:||||:||||:
Db 1201 TCACTCGTGAGCT 1213

RESULT 29
US-09-949-016-14464/C
; Sequence 14464, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14464
; LENGTH: 14205
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14464

Query Match 59.1%; Score 13; DB 4; Length 14205;
Best Local Similarity 76.9%; Pred. No. 92;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 10 CUCGUGAGCUCTT 22
:||||:||||:
Db 9218 CTCGTGAGCTCTT 9206

RESULT 30
US-09-949-016-12247
; Sequence 12247, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12247
; LENGTH: 14395
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12247

Query Match 59.1%; Score 13; DB 4; Length 14395;
Best Local Similarity 76.9%; Pred. No. 92;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AACUCACUCUGA 16
:||||:||||:
Db 11437 AACTCACTCGTGA 11449

Search completed: July 21, 2005, 00:58:19
Job time : 147 secs

81	20	90.9	29736	21	US-10-831-901A-29746	Sequence 29746, A	c 154	15	68.2	1870	9	US-09-864-761-21483	Sequence 21483, A
82	20	90.9	29736	21	US-10-831-901A-29747	Sequence 29747, A	c 155	15	68.2	1966	9	US-09-864-761-21483	Sequence 21483, A
83	20	90.9	29736	21	US-10-831-901A-29749	Sequence 29749, A	c 156	15	68.2	4317	19	US-10-755-889-778	Sequence 778, App
84	20	90.9	29736	21	US-10-831-901A-29750	Sequence 29750, A	c 157	15	68.2	5242	18	US-10-620-914-43	Sequence 43, Appl
85	20	90.9	29736	21	US-10-831-901A-29763	Sequence 29763, A	c 158	14	63.6	20	21	US-10-831-901A-774	Sequence 774, App
86	20	90.9	29736	21	US-10-831-901A-29769	Sequence 29769, A	c 159	14	63.6	20	21	US-10-831-901A-774	Sequence 774, App
87	20	90.9	29736	21	US-10-831-901A-29783	Sequence 29783, A	c 160	14	63.6	348	22	US-10-724-972A-609	Sequence 609, App
88	20	90.9	29736	21	US-10-831-901A-29804	Sequence 29804, A	c 161	14	63.6	535	13	US-10-027-633-187522	Sequence 187522, App
89	20	90.9	29736	21	US-10-831-901A-29805	Sequence 29805, A	c 162	14	63.6	535	17	US-10-027-633-187522	Sequence 187522, App
90	20	90.9	29736	21	US-10-831-901A-29806	Sequence 29806, A	c 163	14	63.6	579	18	US-10-424-593-86013	Sequence 86013, A
91	20	90.9	29736	22	US-10-889-101-9	Sequence 9, Appl	c 163	14	63.6	603	20	US-10-425-115-132098	Sequence 132098, A
92	20	90.9	29736	22	US-10-843-527-238193	Sequence 238193, A	c 164	14	63.6	744	19	US-10-767-701-7756	Sequence 7756, App
93	20	90.9	29738	21	US-10-831-901A-29808	Sequence 29808, A	c 165	14	63.6	932	13	US-10-027-633-170009	Sequence 170009, A
94	20	90.9	29740	21	US-10-831-901A-29810	Sequence 29810, A	c 166	14	63.6	932	13	US-10-027-633-170010	Sequence 170010, A
95	20	90.9	29740	21	US-10-831-901A-29770	Sequence 29770, A	c 167	14	63.6	932	13	US-10-027-633-170011	Sequence 170011, A
96	20	90.9	29740	21	US-10-831-901A-29787	Sequence 29787, A	c 168	14	63.6	932	17	US-10-027-633-170009	Sequence 170009, A
97	20	90.9	29740	21	US-10-831-901A-29799	Sequence 29799, A	c 169	14	63.6	932	17	US-10-027-633-170010	Sequence 170010, A
98	20	90.9	29741	21	US-10-831-901A-29818	Sequence 29818, A	c 170	14	63.6	932	17	US-10-027-633-170010	Sequence 170010, A
99	20	90.9	29741	21	US-10-831-901A-29796	Sequence 29796, A	c 171	14	63.6	1278	17	US-10-027-633-170011	Sequence 170011, A
100	20	90.9	29742	21	US-10-831-901A-29817	Sequence 29817, A	c 172	14	63.6	2950	19	US-10-369-493-43748	Sequence 43748, A
101	20	90.9	29742	21	US-10-839-729-16	Sequence 16, Appl	c 173	14	63.6	3504	17	US-10-437-963-21712	Sequence 21712, A
102	20	90.9	29742	21	US-10-808-187-15	Sequence 15, Appl	c 174	14	63.6	3705	19	US-10-320-797-156	Sequence 156, App
103	20	90.9	29742	21	US-10-808-187-16	Sequence 16, Appl	c 175	14	63.6	3705	19	US-10-437-963-21697	Sequence 21697, A
104	20	90.9	29742	21	US-10-808-187-240	Sequence 240, App	c 176	14	63.6	4314	17	US-10-282-122A-34758	Sequence 34758, A
105	20	90.9	29742	21	US-10-808-187-737	Sequence 737, App	c 177	14	63.6	4704	19	US-10-437-963-21704	Sequence 21704, A
106	20	90.9	29742	21	US-10-808-187-1108	Sequence 1108, App	c 178	14	63.6	4888	19	US-10-437-963-40278	Sequence 40278, A
107	20	90.9	29742	21	US-10-808-187-1590	Sequence 1590, App	c 179	14	63.6	5004	19	US-10-437-963-40272	Sequence 40272, A
108	20	90.9	29742	21	US-10-808-187-1965	Sequence 1965, App	c 180	14	63.6	5388	19	US-10-437-963-40273	Sequence 40273, A
109	20	90.9	29742	21	US-10-889-447-10	Sequence 10, Appl	c 181	14	63.6	7548	19	US-10-437-963-27706	Sequence 27706, A
110	20	90.9	29742	21	US-10-831-901A-29744	Sequence 14, Appl	c 182	14	63.6	8937	19	US-10-437-963-40275	Sequence 40275, A
111	20	90.9	29742	21	US-10-889-101-20	Sequence 10, Appl	c 183	14	63.6	47493	18	US-10-052-482-55	Sequence 55, Appl
112	20	90.9	29742	22	US-10-843-527-238194	Sequence 238194, A	c 184	14	63.6	76180	21	US-10-322-281-432	Sequence 432, App
113	20	90.9	29745	21	US-10-831-901A-29768	Sequence 29768, A	c 185	14	63.6	76698	21	US-10-936-273-30	Sequence 30, Appl
114	20	90.9	29749	21	US-10-831-901A-29811	Sequence 29811, A	c 186	14	63.6	76698	22	US-10-948-947A-1	Sequence 1, Appl
115	20	90.9	29751	21	US-10-831-901A-29809	Sequence 29809, A	c 187	14	63.6	110838	22	US-10-840-512-41	Sequence 41, Appl
116	20	90.9	29751	21	US-10-839-729-14	Sequence 14, Appl	c 188	14	63.6	247544	19	US-10-322-696-55	Sequence 55, Appl
117	20	90.9	29751	21	US-10-856-529-1	Sequence 1, Appl	c 189	13	59.1	20	21	US-10-831-901A-773	Sequence 773, App
118	20	90.9	29751	21	US-10-626-879-67	Sequence 67, Appl	c 190	13	59.1	20	21	US-10-831-901A-787	Sequence 787, App
119	20	90.9	29751	21	US-10-889-447-2	Sequence 2, Appl	c 191	13	59.1	21	17	US-09-842-758-110	Sequence 110, App
120	20	90.9	29751	21	US-10-889-447-2	Sequence 2, Appl	c 192	13	59.1	21	17	US-10-174-333-110	Sequence 110, App
121	20	90.9	29751	21	US-10-699-936-2	Sequence 2, Appl	c 193	13	59.1	21	17	US-10-174-333-186	Sequence 186, App
122	20	90.9	29751	21	US-10-831-901A-29751	Sequence 29751, A	c 194	13	59.1	22	18	US-10-335-977-9995	Sequence 9995, App
123	20	90.9	29751	21	US-10-831-901A-29752	Sequence 29752, A	c 195	13	59.1	25	22	US-10-719-956-73668	Sequence 73668, A
124	20	90.9	29751	21	US-10-831-901A-29778	Sequence 29778, A	c 196	13	59.1	25	22	US-10-719-956-78611	Sequence 78611, A
125	20	90.9	29751	21	US-10-831-901A-29801	Sequence 29801, A	c 197	13	59.1	25	22	US-10-719-956-308817	Sequence 308817, A
126	20	90.9	29751	21	US-10-889-101-2	Sequence 2, Appl	c 198	13	59.1	25	22	US-10-719-956-312434	Sequence 312434, A
127	20	90.9	29751	21	US-10-889-101-2	Sequence 2, Appl	c 199	13	59.1	25	22	US-10-719-956-501537	Sequence 501537, A
128	20	90.9	29751	22	US-10-843-527-238195	Sequence 238195, A	c 200	13	59.1	200	16	US-10-029-386-22454	Sequence 22454, A
129	20	90.9	29754	21	US-10-831-901A-29800	Sequence 29800, A	c 201	13	59.1	275	20	US-10-425-115-18480	Sequence 18480, A
130	20	90.9	29757	21	US-10-831-901A-29769	Sequence 29769, A	c 202	13	59.1	304	21	US-10-487-804-97	Sequence 97, Appl
131	20	90.9	29757	21	US-10-831-901A-29786	Sequence 29786, A	c 203	13	59.1	305	9	US-09-998-598-2109	Sequence 2109, App
132	19	86.4	20	21	US-10-888-401-1	Sequence 1, Appl	c 204	13	59.1	314	20	US-10-425-115-33053	Sequence 33053, A
133	19	86.4	20	21	US-10-831-901A-779	Sequence 779, App	c 205	13	59.1	314	20	US-10-425-115-33053	Sequence 33053, A
134	18	81.8	20	21	US-10-831-901A-781	Sequence 781, App	c 206	13	59.1	329	20	US-10-425-115-72056	Sequence 72056, A
135	18	81.8	20	21	US-10-831-901A-778	Sequence 778, App	c 207	13	59.1	382	20	US-10-425-115-145414	Sequence 145414, A
136	17	77.3	20	21	US-10-831-901A-782	Sequence 782, App	c 208	13	59.1	399	18	US-10-424-599-81718	Sequence 81718, A
137	17	77.3	20	21	US-10-831-901A-777	Sequence 777, App	c 209	13	59.1	399	20	US-10-425-115-168811	Sequence 168811, A
138	17	77.3	20	21	US-10-831-901A-783	Sequence 783, App	c 210	13	59.1	408	18	US-10-424-599-28283	Sequence 28283, A
139	16	72.7	20	21	US-10-831-901A-776	Sequence 776, App	c 211	13	59.1	413	11	US-09-732-627A-4252	Sequence 4252, App
140	16	72.7	20	21	US-10-831-901A-784	Sequence 784, App	c 212	13	59.1	413	11	US-10-242-535A-55663	Sequence 55663, A
141	16	72.7	619	21	US-10-487-901-5047	Sequence 5047, App	c 213	13	59.1	420	18	US-10-085-783A-55663	Sequence 55663, A
142	16	72.7	619	21	US-10-487-901-5767	Sequence 5767, App	c 214	13	59.1	435	13	US-10-027-632-271544	Sequence 271544, A
143	16	72.7	686	21	US-10-487-901-1715	Sequence 1715, App	c 215	13	59.1	435	17	US-10-027-632-271544	Sequence 271544, A
144	16	72.7	686	21	US-10-487-901-5029	Sequence 5029, App	c 216	13	59.1	445	14	US-10-060-036-3310	Sequence 3310, App
145	15	68.2	1891	19	US-10-437-963-101306	Sequence 101306, A	c 217	13	59.1	478	9	US-09-864-761-10972	Sequence 10972, A
146	15	68.2	20	21	US-10-831-901A-775	Sequence 775, App	c 218	13	59.1	486	10	US-09-918-995-7702	Sequence 7702, App
147	15	68.2	20	21	US-10-831-901A-785	Sequence 785, App	c 219	13	59.1	506	18	US-10-424-599-53113	Sequence 53113, A
148	15	68.2	25	22	US-10-719-956-69911	Sequence 69911, A	c 220	13	59.1	506	13	US-10-027-632-315678	Sequence 315678, A
149	15	68.2	25	22	US-10-719-956-69911	Sequence 69911, A	c 221	13	59.1	506	17	US-10-027-633-212565	Sequence 212565, A
150	15	68.2	537	17	US-10-282-122A-38146	Sequence 38146, A	c 222	13	59.1	520	13	US-10-027-633-212565	Sequence 212565, A
151	15	68.2	627	14	US-10-000-256A-81	Sequence 81, Appl	c 223	13	59.1	520	17	US-10-027-633-212565	Sequence 212565, A
152	15	68.2	820	19	US-10-767-701-619	Sequence 619, App	c 224	13	59.1	545	13	US-10-027-632-6883	Sequence 6883, App
153	15	68.2	1101	19	US-10-767-795-2281	Sequence 2281, App	c 225	13	59.1	545	13	US-10-027-632-6884	Sequence 6884, App
			1820	9	US-09-864-761-19662	Sequence 19662, A	c 226	13	59.1	545	17	US-10-027-632-6882	Sequence 6882, App

227	13	59.1	17	US-10-027-632-6883	Sequence 6883, Ap	300	13	59.1	1767	15	US-10-128-714-7568	Sequence 7568, Ap
228	13	545	17	US-10-027-632-6884	Sequence 6884, Ap	ALIGNMENTS						
c 229	13	59.1	16	US-10-029-386-8719	Sequence 8719, Ap							
c 230	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 231	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
232	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
233	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
234	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 235	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 236	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 237	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 238	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 239	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 240	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 241	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 242	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 243	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 244	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 245	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 246	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 247	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 248	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 249	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 250	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 251	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 252	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 253	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 254	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 255	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 256	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 257	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 258	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 259	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 260	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 261	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 262	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 263	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 264	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 265	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 266	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 267	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 268	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 269	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 270	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 271	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 272	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 273	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 274	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 275	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 276	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 277	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 278	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 279	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 280	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 281	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 282	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 283	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 284	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 285	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 286	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 287	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 288	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 289	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 290	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 291	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 292	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 293	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 294	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 295	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 296	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 297	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 298	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 299	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							
c 300	13	59.1	17	US-10-027-632-147918	Sequence 147918, Ap							

RESULT 1

US-10-848-737-1

; Sequence 1, Application US/10848737

; Publication No. US20050004063A1

; GENERAL INFORMATION:

; APPLICANT: HE, MING-LIANG

; APPLICANT: KUNG, HSIANG-FU

; APPLICANT: ZHENG, BOJIAN

; APPLICANT: LIN, MARIE C. M.

; APPLICANT: PENG, YING

; APPLICANT: GUAN, YI

; TITLE OF INVENTION: INHIBITION OF SARS-ASSOCIATED CORONAVIRUS (SCOV)

; FILE REFERENCE: V9661.0080

; CURRENT APPLICATION NUMBER: US/10/848,737

; CURRENT FILING DATE: 2004-05-19

; PRIOR APPLICATION NUMBER: 60/471,901

; PRIOR FILING DATE: 2003-05-19

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: Patent in ver. 3.2

; SEQ ID NO 1

; LENGTH: 22

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE: Description of Combined DNA/RNA Molecule:

; OTHER INFORMATION: Synthetic oligonucleotide SARS1-1

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

; OTHER INFORMATION: oligonucleotide

US-10-848-737-1

Query Match 100.0%; Score 22; DB 21; Length 22;

Best Local Similarity 100.0%; Pred. No. 0.0011;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCGUGAGCUCTT 22

Db 1 GUGAACUCACUCGUGAGCUCTT 22

RESULT 2

US-10-831-901A-780/c

; Sequence 780, Application US/10831901A

; Publication No. US20050100885A1

; GENERAL INFORMATION:

; APPLICANT: Crooke, Stanley T.

; APPLICANT: Ecker, David J.

; APPLICANT: Sampath, Rangarajan

; APPLICANT: Freier, Susan M.

; APPLICANT: Massire, Christian

; APPLICANT: Hofstadler, Steven A.

; APPLICANT: Lowery, Kristin Sannes

; APPLICANT: Swayze, Eric

; APPLICANT: Baker, Brenda F.

; APPLICANT: Bennett, C. Frank

; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe

; PRIOR APPLICATION NUMBER: 60/468,627
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/477,637
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: 60/483,579
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 30063
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 780
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense compound
US-10-831-901A-780

Query Match 90.9%; Score 20; DB 21; Length 20;
Best Local Similarity 75.0%; Pred. No. 0.02;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUGAACUCACUCUGAGCUC 20
|:||||:||||:||||:|
Db 20 GTGAACCTCACTCGTGAGCTC 1

RESULT 3
US-10-755-415-166
; Sequence 166, Application US/10755415
; Publication No. US20050136480A1
; GENERAL INFORMATION:
; APPLICANT: BRAHMACHARI, SAMIR KUMAR
; APPLICANT: DASH, DEBASIS
; APPLICANT: SHARMA, RAMAKANT
; APPLICANT: MAHESHWARI, JITENDRA KUMAR
; TITLE OF INVENTION: A COMPUTER BASED VERSATILE METHOD FOR IDENTIFYING PROTEIN CODING
; FILE REFERENCE: 026033-00029
; CURRENT APPLICATION NUMBER: US/10755,415
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: 10/727,989
; PRIOR FILING DATE: 2003-12-05
; NUMBER OF SEQ ID NOS: 373
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 166
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Sars coronavirus
US-10-755-415-166

Query Match 90.9%; Score 20; DB 22; Length 525;
Best Local Similarity 75.0%; Pred. No. 0.014;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUGAACUCACUCUGAGCUC 20
|:||||:||||:||||:|
Db 76 GTGAACCTCACTCGTGAGCTC 95

RESULT 4
US-10-831-901A-29766
; Sequence 29766, Application US/10831901A
; Publication No. US20050100885A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Ecker, David J.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Freier, Susan M.
; APPLICANT: Massire, Christian
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Lowery, Kristin Sannes
; APPLICANT: Swayze, Eric
; APPLICANT: Baker, Brenda F.
; APPLICANT: Bennett, C. Frank

; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
; FILE REFERENCE: ISIS0083-100 (BIOL00008US)
; CURRENT APPLICATION NUMBER: US/10/831,901A
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/466,426
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: 60/468,562
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/467,770
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 60/468,627
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/477,637
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: 60/483,579
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 30063
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29766
; LENGTH: 1215
; TYPE: DNA
; ORGANISM: Sars Coronavirus
US-10-831-901A-29766

Query Match 90.9%; Score 20; DB 21; Length 1215;
Best Local Similarity 75.0%; Pred. No. 0.013;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUGAACUCACUCUGAGCUC 20
|:||||:||||:||||:|
Db 760 GTGAACCTCACTCGTGAGCTC 779

RESULT 5
US-10-699-936-4
; Sequence 4, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill
; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David E.
; APPLICANT: Jollick, Joseph D.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 1706
; TYPE: DNA
; ORGANISM: SARS-CoV Shanghai LY
US-10-699-936-4

Query Match 90.9%; Score 20; DB 21; Length 1706;
Best Local Similarity 75.0%; Pred. No. 0.012;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUGAACUCACUCUGAGCUC 20
|:||||:||||:||||:|
Db 697 GTGAACCTCACTCGTGAGCTC 716

RESULT 6
US-10-699-936-14
; Sequence 14, Application US/10699936
; Publication No. US20050095582A1
; GENERAL INFORMATION:
; APPLICANT: Gillim-Ross, Laura
; APPLICANT: Taylor, Jill

```

; APPLICANT: Scholl, David R.
; APPLICANT: Wentworth, David B.
; APPLICANT: Jollick, Joseph D.
; TITLE OF INVENTION: Compositions and Methods for Detecting Severe Acute Respiratory
; TITLE OF INVENTION: Syndrome Coronavirus
; FILE REFERENCE: DHI-07986
; CURRENT APPLICATION NUMBER: US/10/699,936
; CURRENT FILING DATE: 2003-11-03
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 1706
; TYPE: DNA
; ORGANISM: SARS coronavirus Shanghai LY
US-10-699-936-14

Query Match          90.9%; Score 20; DB 21; Length 1706;
Best Local Similarity 75.0%; Pred. No. 0.012; Indels 0; Gaps 0;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCGUGAGCUC 20
   |::|::|::|::|::|::|
Db 697 GTGAACCTCACTCGTGAGCTC 716

RESULT 7
US-10-889-447-3
; Sequence 3, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 24774
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ01
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-3

Query Match          90.9%; Score 20; DB 21; Length 24774;
Best Local Similarity 75.0%; Pred. No. 0.009; Indels 0; Gaps 0;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCGUGAGCUC 20
   |::|::|::|::|::|::|
Db 706 GTGAACCTCACTCGTGAGCTC 725

RESULT 8
US-10-831-901A-29748
; Sequence 29748, Application US/10831901A
; Publication No. US20050100885A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Ecker, David J.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Freier, Susan M.
; APPLICANT: Massire, Christian
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Lowery, Kristin Sannes
; APPLICANT: Swayze, Eric
; APPLICANT: Baker, Brenda F.

```

QY 1 GUGAACUCACUCGUGAGCUC 20
|:||||:||||:||||:|
Db 706 GTGAACCTCACTCGTGAGCTC 725

RESULT 10
US-10-889-447-5
; Sequence 5, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BU03
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-5

Query Match 90.9%; Score 20; DB 21; Length 28920;
Best Local Similarity 75.0%; Pred. No. 0.0089;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUGAACUCACUCGUGAGCUC 20
|:||||:||||:||||:|
Db 776 GTGAACCTCACTCGTGAGCTC 795

RESULT 11
US-10-889-447-6
; Sequence 6, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BU04
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-6

Query Match 90.9%; Score 20; DB 21; Length 28920;
Best Local Similarity 75.0%; Pred. No. 0.0089;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUGAACUCACUCGUGAGCUC 20
|:||||:||||:||||:|
Db 776 GTGAACCTCACTCGTGAGCTC 795

RESULT 12
US-10-831-901A-29740
; Sequence 29740, Application US/10831901A
; Publication No. US20050100885A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Ecker, David J.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Freier, Susan M.
; APPLICANT: Massire, Christian
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Lowery, Kristin Sannes
; APPLICANT: Swaze, Eric
; APPLICANT: Baker, Brenda F.
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
; FILE REFERENCE: IS10083-100 (BIOL0008US)
; CURRENT APPLICATION NUMBER: US/10/831,901A
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/466,426
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: 60/468,562
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/467,770
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 60/468,627
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/477,637
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: 60/483,579
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 30063
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29740
; LENGTH: 28920
; TYPE: DNA
; ORGANISM: SARS Coronavirus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7230-7329, 9929-10028, 10137-10236, 11646-11745, 20385-20484,
; LOCATION: 21024-21123, 21753-21852, 22112-22211, 25301-25400
; OTHER INFORMATION: n = A,T,C or G
US-10-831-901A-29740

Query Match 90.9%; Score 20; DB 21; Length 28920;
Best Local Similarity 75.0%; Pred. No. 0.0089;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUGAACUCACUCGUGAGCUC 20
|:||||:||||:||||:|
Db 776 GTGAACCTCACTCGTGAGCTC 795

RESULT 13
US-10-889-101-5
; Sequence 5, Application US/10889101
; Publication No. US20050107324A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Dobie, Kenneth W.
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF CEACAM1 EXPRESSION
; FILE REFERENCE: IS10101-100 (RTS-0655US)
; CURRENT APPLICATION NUMBER: US/10/889,101
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,652
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 28920
; TYPE: DNA

```

; PRIOR APPLICATION NUMBER: 60/468,562
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/467,770
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 60/468,627
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/477,637
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: 60/483,579
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 30063
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29819
; LENGTH: 29013
; TYPE: DNA
; ORGANISM: SARS Coronavirus
US-10-831-901A-29819

Query Match          90.9%; Score 20; DB 21; Length 29013;
Best Local Similarity 75.0%; Pred. No. 0.0089;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1   GUGAACUCACUCGUGAGCUC 20
         |:|::||::||::||::||::||:|
Db       697 GTGAATCACTCGTGAGCTC 716

RESULT 16
US-10-831-901A-29742
; Sequence 29742, Application US/10831901A
; Publication No. US20050100885A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Ecker, David J.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Freier, Susan M.
; APPLICANT: Massire, Christian
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Lowery, Kristin Sannes
; APPLICANT: Swayze, Eric
; APPLICANT: Baker, Brenda F.
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
; TITLE OF INVENTION: Acute Respiratory Syndrome (SARS)
; FILE REFERENCE: ISIS0083-100 (BIOLO0080US)
; CURRENT APPLICATION NUMBER: US/10/831,901A
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/466,426
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: 60/468,562
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/467,770
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 60/468,627
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/477,637
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: 60/483,579
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 30063
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29742
; LENGTH: 29206
; TYPE: DNA
; ORGANISM: SARS Coronavirus
US-10-831-901A-29742

Query Match          90.9%; Score 20; DB 21; Length 29206;
Best Local Similarity 75.0%; Pred. No. 0.0089;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1   GUGAACUCACUCGUGAGCUC 20
         |:|::||::||::||::||::||:|

```

```

Db      751 GTGAACTCACTCGTGAGCTC 770

RESULT 17
US-10-889-447-4
; Sequence 4, Application US/10889447
; Publication No. US20050075307A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF AMINOPEPTIDASE N EXPRESSION
; FILE REFERENCE: RTS-0685US
; CURRENT APPLICATION NUMBER: US/10/889,447
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,670
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 29291
; TYPE: DNA
; ORGANISM: SARS coronavirus isolate BJ02
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n is any nucleotide
US-10-889-447-4

Query Match          90.9%; Score 20; DB 21; Length 29291;
Best Local Similarity 75.0%; Pred. No. 0.0089;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1 GUGAACUCACUCUGAGCUC 20
       |:|::||::||::||::|
Db      726 GTGAACACTCACTCGTGAGCTC 745

RESULT 18
US-10-831-901A-29738
; Sequence 29738, Application US/10831901A
; Publication No. US20050100885A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Ecker, David J.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Freier, Susan M.
; APPLICANT: Massire, Christian A.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Lowery, Kristin Sannes
; APPLICANT: Swayze, Eric
; APPLICANT: Baker, Brenda F.
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
; TITLE OF INVENTION: Acute Respiratory Syndrome (SARS)
; FILE REFERENCE: ISIS0083-100 (BIOL000808US)
; CURRENT APPLICATION NUMBER: US/10/831,901A
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/466,426
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: 60/468,562
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/467,770
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 60/468,627
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/477,637
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: 60/483,579
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 30063
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29738
; LENGTH: 29291

QY      1 GUGAACUCACUCUGAGCUC 20
       |:|::||::||::||::|
Db      726 GTGAACACTCACTCGTGAGCTC 745

RESULT 19
US-10-889-101-4
; Sequence 4, Application US/10889101
; Publication No. US20050107324A1
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Dobie, Kenneth W.
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF CEACAM1 EXPRESSION
; FILE REFERENCE: ISIS0101-100 (RTS-0655US)
; CURRENT APPLICATION NUMBER: US/10/889,101
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,652
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 29291
; TYPE: DNA
; ORGANISM: SARS Coronavirus isolate BJ02
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n = A,T,C or G
US-10-889-101-4

Query Match          90.9%; Score 20; DB 21; Length 29291;
Best Local Similarity 75.0%; Pred. No. 0.0089;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1 GUGAACUCACUCUGAGCUC 20
       |:|::||::||::||::|
Db      726 GTGAACACTCACTCGTGAGCTC 745

RESULT 20
US-10-831-901A-29739
; Sequence 29739, Application US/10831901A
; Publication No. US20050100885A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Ecker, David J.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Freier, Susan M.
; APPLICANT: Massire, Christian A.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Lowery, Kristin Sannes
; APPLICANT: Swayze, Eric
; APPLICANT: Baker, Brenda F.
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
; TITLE OF INVENTION: Acute Respiratory Syndrome (SARS)
; FILE REFERENCE: ISIS0083-100 (BIOL000808US)
; CURRENT APPLICATION NUMBER: US/10/831,901A
; CURRENT FILING DATE: 2004-04-26

```

```

; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Ecker, David J.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Freier, Susan M.
; APPLICANT: Massire, Christian
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Lowery, Kristin Sannes
; APPLICANT: Swayze, Eric
; APPLICANT: Baker, Brenda F.
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
; TITLE OF INVENTION: Acute Respiratory Syndrome (SARS)
; FILE REFERENCE: ISIS0083-100 (BIOL0080US)
; CURRENT APPLICATION NUMBER: US/10/831,901A
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/466,426
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: 60/468,562
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/467,770
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 60/468,627
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/477,637
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: 60/483,579
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 30063
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29741
; LENGTH: 29430
; TYPE: DNA
; ORGANISM: SARS Coronavirus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6464-6563, 11343-11442, 22567-22666
; OTHER INFORMATION: n = A,T,C or G
US-10-831-901A-29741

Query Match          90.9%; Score 20; DB 21; Length 29430;
Best Local Similarity 75.0%; Pred. No. 0.0089;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1  GUGAACUCACUCUGGAGCUC 20
        |:|:|:|:|:|:|:|:|:|:|
Db      760 GTGACTCACTCGTGAGCTC 779

RESULT 23
US-10-889-101-7
; Sequence 7, Application US/10889101
; Publication No. US200501073241
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Dobie, Kenneth W.
; APPLICANT: Jain, Ravi
; TITLE OF INVENTION: MODULATION OF CEACAM1 EXPRESSION
; FILE REFERENCE: ISIS0101-100 (RTS-0655US)
; CURRENT APPLICATION NUMBER: US/10/889,101
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: 60/486,652
; PRIOR FILING DATE: 2003-07-12
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 29430
; TYPE: DNA
; ORGANISM: SARS Coronavirus isolate GZ01
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: n = A,T,C or G

```

US-10-889-101-7

Query Match 90.9%; Score 20; DB 21; Length 29430;
Best Local Similarity 75.0%; Pred. No. 0.0089;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUGAACUCACUCGUGAGCUC 20
|:||||:||||:||||:|
Db 760 GTGAACCTCACTCGTGAGCTC 779

RESULT 24

US-10-831-901A-29802
; Sequence 29802, Application US/10831901A
; Publication No. US20050100885A1

; GENERAL INFORMATION:

; APPLICANT: Crooke, Stanley T.
; APPLICANT: Ecker, David J.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Freier, Susan M.
; APPLICANT: Massire, Christian
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Lowery, Kristin Sannes
; APPLICANT: Swayze, Eric
; APPLICANT: Baker, Brenda F.

; APPLICANT: Bennett, C. Frank

; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe

; FILE REFERENCE: ISI0083-100 (BIOL00008US)

; CURRENT APPLICATION NUMBER: US/10/831,901A

; CURRENT FILING DATE: 2004-04-26

; PRIOR APPLICATION NUMBER: 60/466,426

; PRIOR FILING DATE: 2003-04-28

; PRIOR APPLICATION NUMBER: 60/468,562

; PRIOR FILING DATE: 2003-05-06

; PRIOR APPLICATION NUMBER: 60/467,770

; PRIOR FILING DATE: 2003-04-30

; PRIOR APPLICATION NUMBER: 60/468,627

; PRIOR FILING DATE: 2003-05-06

; PRIOR APPLICATION NUMBER: 60/477,637

; PRIOR FILING DATE: 2003-06-10

; PRIOR APPLICATION NUMBER: 60/483,579

; PRIOR FILING DATE: 2003-06-27

; NUMBER OF SEQ ID NOS: 30063

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 29802

; LENGTH: 29573

; TYPE: DNA

; ORGANISM: SARS Coronavirus

US-10-831-901A-29802

Query Match 90.9%; Score 20; DB 21; Length 29573;
Best Local Similarity 75.0%; Pred. No. 0.0089;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUGAACUCACUCGUGAGCUC 20
|:||||:||||:||||:|
Db 707 GTGAACCTCACTCGTGAGCTC 726

RESULT 25

US-10-831-901A-29803
; Sequence 29803, Application US/10831901A
; Publication No. US20050100885A1

; GENERAL INFORMATION:

; APPLICANT: Crooke, Stanley T.
; APPLICANT: Ecker, David J.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Freier, Susan M.
; APPLICANT: Massire, Christian
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Lowery, Kristin Sannes
; APPLICANT: Swayze, Eric

; APPLICANT: Bennett, C. Frank

; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe

; FILE REFERENCE: ISI0083-100 (BIOL00008US)

; CURRENT APPLICATION NUMBER: US/10/831,901A

; CURRENT FILING DATE: 2004-04-26

; PRIOR APPLICATION NUMBER: 60/466,426

; PRIOR FILING DATE: 2003-04-28

; PRIOR APPLICATION NUMBER: 60/468,562

; PRIOR FILING DATE: 2003-05-06

; PRIOR APPLICATION NUMBER: 60/467,770

; PRIOR FILING DATE: 2003-04-30

; PRIOR APPLICATION NUMBER: 60/468,627

; PRIOR FILING DATE: 2003-05-06

; PRIOR APPLICATION NUMBER: 60/477,637

; PRIOR FILING DATE: 2003-06-10

; PRIOR APPLICATION NUMBER: 60/483,579

; PRIOR FILING DATE: 2003-06-27

; NUMBER OF SEQ ID NOS: 30063

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 29807

; LENGTH: 29573

; TYPE: DNA

; APPLICANT: Baker, Brenda F.
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
; FILE REFERENCE: ISI0083-100 (BIOL00008US)
; CURRENT APPLICATION NUMBER: US/10/831,901A
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/466,426
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: 60/468,562
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/467,770
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 60/468,627
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/477,637
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: 60/483,579
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 30063
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29803
; LENGTH: 29573
; TYPE: DNA
; ORGANISM: SARS Coronavirus
US-10-831-901A-29803

Query Match 90.9%; Score 20; DB 21; Length 29573;
Best Local Similarity 75.0%; Pred. No. 0.0089;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 GUGAACUCACUCGUGAGCUC 20
|:||||:||||:||||:|
Db 707 GTGAACCTCACTCGTGAGCTC 726

RESULT 26

US-10-831-901A-29807
; Sequence 29807, Application US/10831901A
; Publication No. US20050100885A1

; GENERAL INFORMATION:

; APPLICANT: Crooke, Stanley T.
; APPLICANT: Ecker, David J.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Freier, Susan M.
; APPLICANT: Massire, Christian
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Lowery, Kristin Sannes
; APPLICANT: Swayze, Eric
; APPLICANT: Baker, Brenda F.

; APPLICANT: Bennett, C. Frank

; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe

; FILE REFERENCE: ISI0083-100 (BIOL00008US)

; CURRENT APPLICATION NUMBER: US/10/831,901A

; CURRENT FILING DATE: 2004-04-26

; PRIOR APPLICATION NUMBER: 60/466,426

; PRIOR FILING DATE: 2003-04-28

; PRIOR APPLICATION NUMBER: 60/468,562

; PRIOR FILING DATE: 2003-05-06

; PRIOR APPLICATION NUMBER: 60/467,770

; PRIOR FILING DATE: 2003-04-30

; PRIOR APPLICATION NUMBER: 60/468,627

; PRIOR FILING DATE: 2003-05-06

; PRIOR APPLICATION NUMBER: 60/477,637

; PRIOR FILING DATE: 2003-06-10

; PRIOR APPLICATION NUMBER: 60/483,579

; PRIOR FILING DATE: 2003-06-27

; NUMBER OF SEQ ID NOS: 30063

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 29807

; LENGTH: 29573

; TYPE: DNA

```

; APPLICANT: Swayze, Eric
; APPLICANT: Baker, Brenda F.
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
; TITLE OF INVENTION: Acute Respiratory Syndrome (SARS)
; FILE REFERENCE: ISIS0083-100 (BIOL000808)
; CURRENT APPLICATION NUMBER: US/10/831,901A
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/466,426
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: 60/468,562
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/467,770
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 60/468,627
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/477,637
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: 60/483,579
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 30663
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29758
; LENGTH: 29705
; TYPE: DNA
; ORGANISM: SARS Coronavirus
US-10-831-901A-29758

Query Match          90.9%; Score 20; DB 21; Length 29705;
Best Local Similarity 75.0%; Pred: No. 0.0089;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1  GUGAACUCACUCGUGAGCUC 20
      1:|||||:|||||:|||||:
Db      760  GTGAACCTCACTCGTGAGCTC 779

RESULT 29
US-10-831-901A-29791
; Sequence 29791, Application US/10831901A
; Publication No. US20050100885A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Ecker, David J.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Freier, Susan M.
; APPLICANT: Massire, Christian
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Lowery, Kristin Sannes
; APPLICANT: Swayze, Eric
; APPLICANT: Baker, Brenda F.
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
; TITLE OF INVENTION: Acute Respiratory Syndrome (SARS)
; FILE REFERENCE: ISIS0083-100 (BIOL000808)
; CURRENT APPLICATION NUMBER: US/10/831,901A
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/466,426
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: 60/468,562
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/467,770
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 60/468,627
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/477,637
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: 60/483,579
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 30663
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29791
; LENGTH: 29705

```

```

; TYPE: DNA
; ORGANISM: SARS Coronavirus
US-10-831-901A-29791

Query Match          90.9%; Score 20; DB 21; Length 29705;
Best Local Similarity 75.0%; Pred. No. 0.0089;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1 GUGAACUCACUCGUGAGCUC 20
Db      760 GTGAACCTCACTCGTGAGCTC 779

RESULT 30
US-10-831-901A-29756
; Sequence 29756, Application US/10831901A
; Publication No. US20050100895A1
; GENERAL INFORMATION:
; APPLICANT: Crooke, Stanley T.
; APPLICANT: Ecker, David J.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Freier, Susan M.
; APPLICANT: Massire, Christian
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Lowery, Kristin Sannes
; APPLICANT: Swayze, Eric
; APPLICANT: Baker, Brenda F.
; APPLICANT: Bennett, C. Frank
; TITLE OF INVENTION: Compositions And Methods For The Treatment Of Severe
; TITLE OF INVENTION: Acute Respiratory Syndrome (SARS)
; FILE REFERENCE: ISIS0083-100 (BIOL00008US)
; CURRENT APPLICATION NUMBER: US/10/831,901A
; CURRENT FILING DATE: 2004-04-26
; PRIOR APPLICATION NUMBER: 60/466,426
; PRIOR FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: 60/468,562
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/467,770
; PRIOR FILING DATE: 2003-04-30
; PRIOR APPLICATION NUMBER: 60/468,627
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: 60/477,637
; PRIOR FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: 60/483,579
; PRIOR FILING DATE: 2003-06-27
; NUMBER OF SEQ ID NOS: 30063
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29756
; LENGTH: 29706
; TYPE: DNA
; ORGANISM: SARS Coronavirus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 20688
; OTHER INFORMATION: n = A,T,C or G
US-10-831-901A-29756

Query Match          90.9%; Score 20; DB 21; Length 29706;
Best Local Similarity 75.0%; Pred. No. 0.0089;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1 GUGAACUCACUCGUGAGCUC 20
Db      760 GTGAACCTCACTCGTGAGCTC 779

Search completed: July 21, 2005, 01:08:21
Job time : 609 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2005, 18:32:31 ; Search time 1685 Seconds
(without alignments)
632.650 Million cell updates/sec

Title: US-10-848-737-1

Perfect score: 22

Sequence: 1 gugaacucacugagucguctt 22

Scoring table:

OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 300 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hlg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sta.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	90.9	1706	14	AY322205 SARS CORO
2	20	90.9	14011	14	AY534758 SARS CORO
3	20	90.9	26333	14	AY286320 SARS CORO
4	20	90.9	29013	14	AY463060 SARS CORO
5	20	90.9	29350	14	AY394999 SARS CORO
6	20	90.9	29350	14	AY395000 SARS CORO
7	20	90.9	29350	14	AY395001 SARS CORO
8	20	90.9	29350	14	AY395002 SARS CORO
9	20	90.9	29433	14	AY394977 SARS CORO
10	20	90.9	29530	14	AY394985 SARS CORO
11	20	90.9	29573	14	AY394981 SARS CORO
12	20	90.9	29573	14	AY394982 SARS CORO
13	20	90.9	29573	14	AY394983 SARS CORO
14	20	90.9	29577	14	AY559094 SARS CORO
15	20	90.9	29592	14	AY463059 SARS CORO
16	20	90.9	29620	14	AY395004 SARS CORO
17	20	90.9	29640	14	AY394978 SARS CORO
18	20	90.9	29645	14	AY394979 SARS CORO
19	20	90.9	29646	14	AY394982 SARS CORO

20	20	90.9	29647	14	AY395003 SARS CORO
21	20	90.9	29661	14	AY559086 SARS CORO
22	20	90.9	29665	14	AY394988 SARS CORO
23	20	90.9	29670	14	AY559082 SARS CORO
24	20	90.9	29683	14	AY394996 SARS CORO
25	20	90.9	29683	14	AY394997 SARS CORO
26	20	90.9	29699	14	AY394983 SARS CORO
27	20	90.9	29705	14	AY283795 SARS CORO
28	20	90.9	29705	14	AY394980 SARS CORO
29	20	90.9	29706	14	AY283797 SARS CORO
30	20	90.9	29709	14	AY394987 SARS CORO
31	20	90.9	29710	14	AY559091 SARS CORO
32	20	90.9	29711	14	AY283794 SARS CORO
33	20	90.9	29711	14	AY283796 SARS CORO
34	20	90.9	29711	14	AY283798 SARS CORO
35	20	90.9	29711	14	AY427439 SARS CORO
36	20	90.9	29712	14	AY559093 SARS CORO
37	20	90.9	29713	14	AY559085 SARS CORO
38	20	90.9	29713	14	AY559092 SARS CORO
39	20	90.9	29714	14	AY559088 SARS CORO
40	20	90.9	29715	14	AY297028 SARS CORO
41	20	90.9	29715	14	AY461660 SARS CORO
42	20	90.9	29715	14	AY559097 SARS CORO
43	20	90.9	29716	14	AY559081 SARS CORO
44	20	90.9	29716	14	AY559087 SARS CORO
45	20	90.9	29716	14	AY595412 SARS CORO
46	20	90.9	29717	14	AY559089 SARS CORO
47	20	90.9	29719	14	AY559095 SARS CORO
48	20	90.9	29720	14	AY304495 SARS CORO
49	20	90.9	29720	14	AY559096 SARS CORO
50	20	90.9	29722	14	AY559084 SARS CORO
51	20	90.9	29725	14	AP006558 SARS CORO
52	20	90.9	29725	14	AY278488 SARS CORO
53	20	90.9	29725	14	AY321118 SARS CORO
54	20	90.9	29725	14	AY559090 SARS CORO
55	20	90.9	29726	14	AY394981 SARS CORO
56	20	90.9	29727	14	AP006557 SARS CORO
57	20	90.9	29727	14	AP006559 SARS CORO
58	20	90.9	29727	14	AP006560 SARS CORO
59	20	90.9	29727	14	AP006561 SARS CORO
60	20	90.9	29727	14	AY278741 SARS CORO
61	20	90.9	29727	14	AY291315 SARS CORO
62	20	90.9	29727	14	AY362698 SARS CORO
63	20	90.9	29727	14	AY362699 SARS CORO
64	20	90.9	29727	14	AY502924 SARS CORO
65	20	90.9	29727	14	AY714217 SARS CORO
66	20	90.9	29728	14	AY394850 SARS CORO
67	20	90.9	29729	14	AY291451 SARS CORO
68	20	90.9	29729	14	AY394986 SARS CORO
69	20	90.9	29729	14	AY502923 SARS CORO
70	20	90.9	29729	14	AY502925 SARS CORO
71	20	90.9	29729	14	AY502926 SARS CORO
72	20	90.9	29729	14	AY502927 SARS CORO
73	20	90.9	29729	14	AY502928 SARS CORO
74	20	90.9	29729	14	AY502929 SARS CORO
75	20	90.9	29729	14	AY502930 SARS CORO
76	20	90.9	29729	14	AY502931 SARS CORO
77	20	90.9	29729	14	AY502932 SARS CORO
78	20	90.9	29731	14	AY304488 SARS CORO
79	20	90.9	29732	14	AY279354 SARS CORO
80	20	90.9	29732	14	AY508724 SARS CORO
81	20	90.9	29733	14	AY278554 SARS CORO
82	20	90.9	29736	14	AY278554 SARS CORO
83	20	90.9	29736	14	AY282752 SARS CORO
84	20	90.9	29736	14	AY345986 SARS CORO
85	20	90.9	29736	14	AY345987 SARS CORO
86	20	90.9	29736	14	AY345988 SARS CORO
87	20	90.9	29736	14	AY394989 SARS CORO
88	20	90.9	29736	14	AY394990 SARS CORO
89	20	90.9	29736	14	AY394991 SARS CORO
90	20	90.9	29736	14	AY394992 SARS CORO
91	20	90.9	29736	14	AY394993 SARS CORO
92	20	90.9	29736	14	AY394998 SARS CORO

93	20	90.9	29738	14	AY350750	AY350750 SARS coro	c 166	15	68.2	12050	10	MMU575748	AJ575748 Mus muscu
94	20	90.9	29738	14	AY357075	AY357075 SARS coro	167	15	68.2	43599	9	HSU15177	U15177 Human cosmi
95	20	90.9	29740	14	AY278490	AY278490 SARS coro	168	15	68.2	50389	1	AE014137	AE014137 Streptoco
96	20	90.9	29740	14	AY310120	AY310120 SARS coro	c 169	15	68.2	51430	2	AC017981	AC017981 Drosophil
97	20	90.9	29740	14	AY485278	AY485278 SARS coro	c 170	15	68.2	54668	6	AX602203	AX602203 Sequence
98	20	90.9	29741	14	AY304486	AY304486 SARS coro	c 171	15	68.2	61524	9	AL353725	AL353725 Human DNA
99	20	90.9	29741	14	AY485277	AY485277 SARS coro	c 172	15	68.2	81786	9	AC003002	AC003002 Human DNA
100	20	90.9	29742	14	AY278491	AY278491 SARS coro	c 173	15	68.2	83369	8	ATT30A10	AL117386 Arabidops
101	20	90.9	29745	14	AY278487	AY278487 SARS coro	174	15	68.2	98432	8	AC007187	AC007187 Arabidops
102	20	90.9	29745	14	AY357076	AY357076 SARS coro	175	15	68.2	110000	1	CP000003_01	Continuation (2 of
103	20	90.9	29745	14	AY654624	AY654624 SARS coro	176	15	68.2	110000	2	AC110967_2	Continuation (3 of
104	20	90.9	29749	14	AY351680	AY351680 SARS coro	177	15	68.2	140012	8	AC135500	AL765500 Oryza sat
105	20	90.9	29751	14	AY274119	AY274119 SARS coro	178	15	68.2	143050	1	SAG766843	AL766843 Streptoco
106	20	90.9	29751	14	AY323977	AY323977 SARS coro	c 179	15	68.2	147494	2	AC015740	AC015740 Homo sapi
107	20	90.9	29754	14	AY313906	AY313906 SARS coro	c 180	15	68.2	149515	2	CR847787	CR847787 Danio rer
108	20	90.9	29757	14	AY278489	AY278489 SARS coro	181	15	68.2	155535	9	AC016909	AC016909 Homo sapi
109	20	90.9	29760	14	AY390556	AY390556 SARS coro	182	15	68.2	157165	4	AC087843	AC087843 Bos tauru
110	20	90.9	29765	14	AY394994	AY394994 SARS coro	183	15	68.2	157662	9	AL139035	AL139035 Human DNA
111	20	90.9	29765	14	AY394995	AY394995 SARS coro	c 184	15	68.2	158692	2	CR356245	CR356245 Danio rer
112	20	90.9	29767	14	AY559083	AY559083 SARS coro	c 185	15	68.2	160950	2	AC011849	AC011849 Homo sapi
113	17	77.3	147433	2	AC137257	AC137257 Rattus no	186	15	68.2	162109	5	BX001032	BX001032 Zebrafish
114	17	77.3	156584	2	AC120103	AC120103 Rattus no	c 187	15	68.2	163242	2	AC132755	AC132755 Rattus no
115	17	77.3	221722	2	AC130567	AC130567 Rattus no	c 188	15	68.2	163263	2	AC121352	AC121352 Rattus no
116	16	72.7	838	8	AK062203	AK062203 Oryza sat	c 189	15	68.2	164000	9	AL591682	AL591682 Human DNA
117	16	72.7	947	8	AK099574	AK099574 Oryza sat	c 190	15	68.2	164766	10	AC144792	AC144792 Mus muscu
118	16	72.7	120551	9	AL391215	AL391215 Human DNA	c 191	15	68.2	174839	9	AL157395	AL157395 Human DNA
119	16	72.7	127267	8	CNS0984U	EX119951 Oryza sat	c 192	15	68.2	184783	8	AC136972	AC136972 Oryza sat
120	16	72.7	158563	2	AC150484	AC150484 Bos tauru	c 193	15	68.2	188901	2	AC132886	AC132886 Mus muscu
121	16	72.7	176629	2	AC055792	AC055792 Homo sapi	194	15	68.2	189318	2	CR376821	CR376821 Danio rer
122	16	72.7	218200	2	AC141165	AC141165 Rattus no	195	15	68.2	189757	3	AC008186	AC008186 Drosophil
123	16	72.7	227026	2	AC136659	AC136659 Rattus no	c 196	15	68.2	190708	9	AC005261	AC005261 Homo sapi
124	16	72.7	237554	2	AC114189	AC114189 Rattus no	c 197	15	68.2	191928	10	AC138714	AC138714 Mus muscu
125	16	72.7	238390	2	AC117035	AC117035 Rattus no	c 198	15	68.2	198116	10	AC130217	AC130217 Mus muscu
126	16	72.7	238394	2	AC131227	AC131227 Rattus no	c 199	15	68.2	199754	8	ATCHRIV26	AL161514 Arabidops
127	16	72.7	247089	2	AC150695	AC150695 Bos tauru	c 200	15	68.2	200702	8	AC130278	AC130278 Mus muscu
128	15	68.2	534	6	CQ649388	CQ649388 Sequence	c 201	15	68.2	204168	10	AC147989	AC147989 Mus muscu
129	15	68.2	537	6	AX607323	AX607323 Sequence	c 202	15	68.2	206615	2	AC139041	AC139041 Mus muscu
130	15	68.2	627	6	AX593704	AX593704 Sequence	c 203	15	68.2	206967	10	AC135292	AC135292 Mus muscu
131	15	68.2	1009	10	BC018219	BC018219 Rattus no	c 204	15	68.2	210344	9	AC007100	AC007100 Homo sapi
132	15	68.2	1794	10	BC018219	BC018219 Rattus no	c 205	15	68.2	219730	2	AC110957	AC110957 Rattus no
133	15	68.2	1820	6	CQ057040	CQ057040 Sequence	c 206	15	68.2	221753	2	AC114835	AC114835 Rattus no
134	15	68.2	1820	6	CQ076301	CQ076301 Sequence	c 207	15	68.2	227774	2	AC123156	AC123156 Rattus no
135	15	68.2	1820	6	CQ107294	CQ107294 Sequence	c 208	15	68.2	229061	9	AC007786	AC007786 Homo sapi
136	15	68.2	1820	6	CQ181412	CQ181412 Sequence	c 209	15	68.2	230254	2	AC074163	AC074163 Mus muscu
137	15	68.2	1820	6	CQ205753	CQ205753 Sequence	c 210	15	68.2	232401	2	AC106137	AC106137 Rattus no
138	15	68.2	1820	6	CQ267291	CQ267291 Sequence	c 211	15	68.2	233788	2	AC108582	AC108582 Rattus no
139	15	68.2	1820	6	CQ304315	CQ304315 Sequence	c 212	15	68.2	234017	2	AC126997	AC126997 Rattus no
140	15	68.2	1820	6	CQ341583	CQ341583 Sequence	c 213	15	68.2	235382	2	AC130104	AC130104 Rattus no
141	15	68.2	1870	6	CQ078257	CQ078257 Sequence	c 214	15	68.2	243313	3	AE003834	AE003834 Drosophil
142	15	68.2	1870	6	CQ109247	CQ109247 Sequence	c 215	15	68.2	245066	2	AC094644	AC094644 Rattus no
143	15	68.2	1870	6	CQ147884	CQ147884 Sequence	c 216	15	68.2	252411	2	AC094615	AC094615 Rattus no
144	15	68.2	1870	6	CQ183233	CQ183233 Sequence	c 217	15	68.2	256781	2	AC097952	AC097952 Rattus no
145	15	68.2	1870	6	CQ207657	CQ207657 Sequence	c 218	15	68.2	260227	2	AC150660	AC150660 Mus muscu
146	15	68.2	1870	6	CQ269291	CQ269291 Sequence	c 219	15	68.2	261031	2	AC140768	AC140768 Rattus no
147	15	68.2	1870	6	CQ306316	CQ306316 Sequence	c 220	15	68.2	270221	2	AC107500	AC107500 Rattus no
148	15	68.2	1870	6	CQ343438	CQ343438 Sequence	c 221	15	68.2	272796	2	AC133856	AC133856 Rattus no
149	15	68.2	1966	6	CQ052020	CQ052020 Sequence	c 222	15	68.2	311600	1	AP005141	AP005141 Streptoco
150	15	68.2	1966	6	CQ067063	CQ067063 Sequence	223	14	63.6	428	6	AR504034	AR504034 Sequence
151	15	68.2	1966	6	CQ094118	CQ094118 Sequence	224	14	63.6	428	6	AR519316	AR519316 Sequence
152	15	68.2	1966	6	CQ171485	CQ171485 Sequence	c 225	14	63.6	488	12	AY200782	AY200782 Arabidops
153	15	68.2	1966	6	CQ200599	CQ200599 Sequence	c 226	14	63.6	534	6	CQ649386	CQ649386 Sequence
154	15	68.2	1966	6	CQ254675	CQ254675 Sequence	c 227	14	63.6	590	11	HS635N10S	AL035144 H. sapiens
155	15	68.2	1966	6	CQ291760	CQ291760 Sequence	c 228	14	63.6	788	11	BV032952	BV032952 S212P6576
156	15	68.2	1966	6	CQ328817	CQ328817 Sequence	c 229	14	63.6	819	11	BV014760	BV014760 S212P6849
157	15	68.2	3066	8	AK110007	AK110007 Oryza sat	230	14	63.6	1081	9	HSJ32831S	AJ32831S Homo sapi
158	15	68.2	4064	6	CQ572246	CQ572246 Sequence	c 231	14	63.6	1142	9	HSJ332579	AJ332579 Homo sapi
159	15	68.2	4064	6	CQ574070	CQ574070 Sequence	c 232	14	63.6	1143	3	AF083649	AF083649 Caenorhab
160	15	68.2	4089	6	CQ576314	CQ576314 Sequence	c 233	14	63.6	1341	6	AX061735	AX061735 Sequence
161	15	68.2	4317	9	HSJ276316	AJ276316 Homo sapi	c 234	14	63.6	1341	8	AF058905	AF058905 Oryza sat
162	15	68.2	8842	2	AC135136	AC135136 Rattus no	c 235	14	63.6	1347	3	BT008754	BT008754 Arabidops
163	15	68.2	11518	10	MMU511265	AJ511265 Mus muscu	c 236	14	63.6	1497	3	AF083644	AF083644 Caenorhab
164	15	68.2	11570	1	AE009959	AE009959 Streptoco	c 237	14	63.6	1514	8	AY099805	AY099805 Arabidops
165	15	68.2	12028	1	AE006478	AE006478 Streptoco	c 238	14	63.6	1539	3	AF083643	AF083643 Caenorhab

Ng,L.C., Se-Thoe,S.Y., Oon,L., Bai,X., Stanton,L.W., Ruan,Y.,
Miller,L.D., Vega,V.B., James,L., Ooi,P.L., Kai,C.S., Olsen,S.J.,
Ang,B. and Leo,Y.S.
TITLE
Laboratory-acquired severe acute respiratory syndrome (SARS) -
Singapore 2003
JOURNAL
REFERENCE
AUTHORS
2 (bases 1 to 14011)
Wei,C.L., Lee,C., Lin,S., Thoreau,H., Vega,V.B., Stanton,L.W. and
Ruan,Y.
TITLE
Direct Submission
JOURNAL
SUBMITTED (28-JAN-2004) Genome Institute of Singapore, 60, Biopolis
Street, 02-01, Genome, Singapore 138672, Singapore
FEATURES
Location/Qualifiers
1. .14011
/organism="SARS coronavirus Sin0409"
/mol_type="genomic RNA"
/db_xref="taxon:266147"
/country="Singapore"
ORIGIN
Query Match 90.9%; Score 20; DB 14; Length 14011;
Best Local Similarity 75.0%; Pred.No. 0.013;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GUGAACUCACUCUGAGCUC 20
|:||||:||||:||||:
Db 231 GTGAACCTCACTCGAGCTC 250
RESULT 3
AY286320
LOCUS
DEFINITION
SARS coronavirus ZJ01, partial genome.
ACCESSION
AY286320
VERSION
AY286320.4 GI:39980888
KEYWORDS
SARS coronavirus ZJ01
ORGANISM
SARS coronavirus ZJ01
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE
1 (bases 1 to 26333)
Li,L., Wang,Z., Lu,Y., Bao,Q., Chen,S., Wu,N., Cheng,S., Weng,J.,
Zhang,Y., Yan,J., Mei,L., Wang,X., Zhu,H., Yu,Y., Zhang,M., Li,M.,
Yao,J., Lu,Q., Yao,P., Bo,X., Mo,J., Wang,S. and Hu,S.
TITLE
Severe acute respiratory syndrome-associated coronavirus genotype
and its characterization
JOURNAL
MEDLINE
22889812
PUBMED
14527350
REFERENCE
2 (bases 1 to 26333)
Wang,Z.G., Li,L.J., Luo,Y., Zhang,J.Y., Wang,M.Y., Cheng,S.Y.,
Zhang,Y.J., Wang,X.M., Lu,Y.Y., Wu,N.P., Mei,L.L., and Wang,Z.X.
TITLE
Molecular biological analysis of genotyping and phylogeny of severe
acute respiratory syndrome associated coronavirus
JOURNAL
MEDLINE
14733771
PUBMED
14733771
REFERENCE
3 (bases 1 to 26333)
Wang,Z., Cheng,S. and Zhang,Y.
TITLE
Direct Submission
JOURNAL
SUBMITTED (28-APR-2003) Department of Microbiology, Zhejiang CDC,
17 Laozhedazi Rd., Hangzhou, Zhejiang 310009, China
REFERENCE
4 (bases 1 to 26333)
Wang,Z., Cheng,S. and Zhang,Y.
TITLE
Direct Submission
JOURNAL
SUBMITTED (19-JUN-2003) Department of Microbiology, Zhejiang CDC,
17 Laozhedazi Rd., Hangzhou, Zhejiang 310009, China
REMARK
Sequence update by submitter
5 (bases 1 to 26333)
Wang,Z., Cheng,S. and Zhang,Y.
TITLE
Direct Submission
JOURNAL
SUBMITTED (13-NOV-2003) Department of Microbiology, Zhejiang CDC,
17 Laozhedazi Rd., Hangzhou, Zhejiang 310009, China
REMARK
Sequence update by submitter

REFERENCE
AUTHORS
TITLE
JOURNAL
SUBMITTED (17-DEC-2003) Department of Microbiology, Zhejiang CDC,
17 Laozhedazi Rd., Hangzhou, Zhejiang 310009, China
REMARK
Sequence update by submitter
COMMENT
On Dec 17, 2003 this sequence version replaced gi:38304880.
FEATURES
Location/Qualifiers
1. .26333
source

/organism="SARS coronavirus ZJ01"
/mol_type="genomic RNA"
/isolate="ZJ01b"
/isolation_source="cultured virus in Vero cell inoculated
with throat swab from patient with severe acute
respiratory syndrome (SARS)"
/db_xref="taxon:230471"
/country="China: Hangzhou"
/note="Isolated April 21, 2003"
1. .248
5'UTR
misc_feature
51. .56
/note="transcription regulatory sequence leader; TRS"
join(249..13382,13382..21469)
/notes="ribosomal slippage"
/codon_start=1
/product="orf1ab polyprotein"
/protein_id="AAP49011.4"
/db_xref="GI:38324306"
/translation="MESLVLGVNEKTHVLSPLVLOVRDVLVRGFGDSVEEALSEARE
HLKNGTCGLVELEKGVLPQLEQYVFIKESDALSNNHKKVLEVLVAEMDGIQVGRSGI
TLGLVPHVGEPIATRNLLRNKNGKAGGHSYGIDLSKSYDLGDLGDDPDEDYEQN
WNTKSGSALRELTRELNGAVTRYVDNNFCGDPGLDCIKDFLARAGKMCFTLSEQ
LDYISKRVCYCRDHEIAWFTERSKSYEQHTPEIKSAKDFTFGKSCPKFVFP
LANSKVKVIQPRVEKKTEGFMGRISRVYPVSPQECNNMHLSTLMKCNHSECVSWOTC
DELKATCEHGTENLIEGPTTCGYLPTNAVVKPCACQDPEIGHSVADVHNHNS
IETRLKGRTRCFGCGFAYVGCYNKRAYWPRASADIGSGHTGTGNVETNEDL
LELSKRVNINIVGDFHLENEVAIIILASFSASTAFIDTIKSLDKSKFTIVESCQN
YKVTGKPVKGANVQQRSVITPLCGFSPQAGVIRISPARTLDANHSIPOLQRAA
VITLDIGSRQLSDVADMYTTLSDTNSVIIMAYVTGLVQQTQSWLSNLLGTTVEKL
RPIFWEAKLSAGVEFLKADWEILKITVDFIVKQGIQVANSNDKCVKCFIDV
NKALEMCIQVTTIAGAKLSNLGEVFIQSKGLYRCQIKRKEQLLAPLKPKEVT
FLEGSHDPLTSEEVLKNGLEALETPVDSFNGAIVGTVCVNGLLMLLETKEQ
YCALSPGLATNNVRLKAGAPIKGVTFGEDTWVEQGVKNRITFELDERVDKVLNE
KCSVTVESGTEVTFACVAAAVKTLQFVSLDLTNMGLDDEWSVATFLPDDAGE
ENFSRMYCSFYPPDEEEDDAEIEEDTCEHEYGTEDDYQGLPFGFSGASTVR
NEEEDDLDDTTQSEIPEPEPTPEPVPQFTGYLKTNDNVAIKCDVIVKEAOSA
NPWVIANAIIHLKGGGVAGALNKATNGAMOKESDDYIKLNGPLTVGSGCLISGHNL
AKCLHVGNLACEDIOLKAAVENSQDILLAPLSAGI FGAKPLQSLQVCVQT
VITQVYIANDKALYEQVMDYLDNLKPRVAPKQEPNTEDESKTEESVVKQPVQV
KPKIKACIDEVITTEETFTLNKLLFADINGKLYHDSQNMURGDSMFLEKADPYM
VGDVITSGDITCVIPIPKKAGGTEMLSRALKKVPVDEYITTPYGGCGAGYTLLEAKT
ALKCKSAFYVLPSEAPNAKEEILGTVSNLREMLAHAEETRLKMPICHDVRAIMATI
QRKYGIKIQEGIVDGVRRFFVTSKEPVASITKLNSNEPLVTMPIGYVTHGFNLE
EACWRSIKAPAVSVSPDATTYNGYLTSSSTSEEHFVETVSLAGSYRDSVSG
QTELGVEFLKGDKLVHYHTLSPVEPHLDGVELDLKLSLLEVEKTIKVFTTVD
NYNHTQLVMSMTYQQQPGFYLDGADVTKIKPHVNHGKTF FVPSDDTLKSAFE
VITLHDESFLGRYMSALNHTKKWFPQVGLTIS KWADNNCYLSLLAQQLQELFEN
APALQAYYRARGAANFCAI LAYSNKTVEGLDRETMTLLHQLANLSAKRVLN
VYCKHCGQXTTLTGVEAVMGTLSYDNLKTVGSIPCVCGRDATQYLPFGQSSFFVM
SAPPKYLQOQTFILCANVTGYOQHVTHTAKETLYRIDGAKHTKMSYKGPVTD
VFYKETSITTIKPVSLKDGVTYIEIPKLGQYKKNAYTEQIDILVPTQPLNA
SFENFKLTCSTNFKPADLNMOTGFTKPSARELSVTFPDLNGDVDAIDRHSYASFK
KAGLHKPTVWHINQATTIKTTPKNTWCRLCRLMSTKPVDSNSFVLADELGHEDLMA
ACESQOQTESEVENPTIQKEVIECDVKTTEVGVNVLKPSDEGVKTVPTQGLHEDLMA
AVENTSITIKKPNELSLALGKTIATHGIAINSVPWSKILAYVFPFLGQAATITSN
GINYVSKPESKFLTIAMWLLLSICLSLICTVTAAGVLLSNFSGVCGVRELYL
NNSNVTMDFCGSGFPCSICLSGLDSDSPALETTIQVTISSYKLDLTILGLAEVWL
CAKRLAQRVNNYMPVFTLLFQLCFTFKTSNRIASLPTTIKAKNSVKSACLCLDA
GINYVSKPESKFLTIAMWLLLSICLSLICTVTAAGVLLSNFSGVCGVRELYL
AYMLFTKFFYLGLSALMQVFFGYFASHFISNSLWMLFIISIVQMAPAVSAMRYIFF
ASFYIYIKSVYHIMDGTSTCMYCRNRATRECTIVNGMKRSFYVYVAGKAGLHL
KTHWNLCDTCTGCTSTFIDEVARDLSQFKRPINPTDQSSYIVDSVAVKNGALHL
YFDKAGQKTYERHPLSHFVNLDNRANNTKSLPINIVIVEDGSKCKDESKASVYV
SQLMCPILLLOALVSDSTSEVSMFDAYVDFSATFSVPMEKALKAVATASE
LAKGVALDGLVSTFSAARQGVVDVTDVKDIECLKSHHSDLEVTGSCNNFMITY

KNVENMTPLDGLACIDCNARHINQVAKSHNVSILINWVDYMSLSBOLRQKQIRSAKK
NIIIPFRUCATITROIVNVIITKISLKGKIVSTCFKMLKATILCLVLAALVCYIVMPV
HTISHDGYNEIBKQIAGVTRDIISTDDCFANKAGDAWDFQSGGSGYNDKSC
PVAAAITRIGETPLGPTGLVIRANGDFELHFLPRVFSAGNICYTPPSKLIBESFOA
TSACVLAECTIPKDMGKVPVYCYDITLLEGGISYSELRPDTRVYMLDGSIIQFNT
YLEGSVRVTTFAEYCHCTCERSEVIGICLSTSGRWLNNEHYRALSGVFCGVDAW
LIANITFPVQPCALDVASVANGGIIAILVTCAYYFMKPRVRGSEYHNVVAANAL
LFLMSFTILCPAISPVGYSYFYDILYTFDNDVSFLAHQWFAFSPVFPFMT
AIYVFCISLKHCHWFNNYLRKRMVNGVTFSTFEAAALCNDFSNKEMLYLKRSETLL
PLQYNYRLALYNYKYFSGALDITTSYREAAACHLAALNDFNSGADVLVYQPOTSI
TSVLAQSGFKMAFPSPKQVGVCMQVOTCGTITLNGMLDVTVCPRHVICTARDMLNP
NYEDLLRKSNSHSLFOAGNQLRVIGHSNOCILRLKVDTSNPKTPKPKFYKRVIRQPG
TFSVLACVNGSPSGVQCAMRPHITKGSFNGSCSGVGNIDYCVSFCYMHMELP
TVHAGVDSGKFPFVPTQAAAGTDTITILNLAWLYAAVINGDWFNLNFTT
LNDNFVAMKYNEPFLQDHDVITIGLPSAQGTAVLDMCAALKEGLQNGNRTILGS
TILDEFTPTDVGSGVTFQGFKKIVKGTTHMMLLTFTSLTILVLSQWTSQSLFF
VIENALPFTLGMIAIAACAMLAVKHKAFLCLFLPLSLATVAVFNNVWPSAWMRI
MTWLELADTSLGVRKDXCVMAISALVLLILMTARTVYDAAARVMTLMNVITLVYKV
YGNALDOALSMALVLSVTSNYSGVVTTIMFLARALVFCVRYEYLLFITGNLQCI
MLVYFGYCCCCFGFLCILLNRYFTLGVYDYLSTQBFYRNMGGLLPPKSSIDA
FKNLIKLLGGKPKCIKATVQSKMSDVKTSVLLSVLQOLRVESSSKLMAQCQLH
NDILAKDTTEAPEKMWVLSVLLSMQGAVIDINRLCEMLDNRAITQAIASEFSLPS
YAAVATAQEAAYEQAVANGSEVVLKLLKSLNAKSEBDRDAAMORKLEKWAQAMTO
MYQABSEDKRAVTSAMQTMFLMLKLENDALNNIINNARDGCPNLIPITTAAK
LMVVPDYGTYKNTDGNITFYASALWEIQOVVDAUSKHVQLSEINMDSNIPITTAAK
VTALRANSVYKLQNELSPVALRQMSCAAGTTQACTDDNALAYNNNSKGRFVLLALL
SDHQLKWARFPVPSKDGSTGTYTELEPPCFRVTPTKPGKVKLYIFYIKGLNLRNGVLL
GSLAATVRLQGNATVPAHSTVLSFCAPAVPAKAYDYLASGQGPITNCVXMLCTH
TGTOGATVTPENMDOESFGGASCLYCHDIDHPNPKGCDLKGKYYOIPITTCAD
PVGFTLRNTVCTVGMKGYGCSDDQLREFMQSADASTFLNRVCGVSAARLTPCGTG
TSDVYVRAFDINERKAVGAFKLTNCCPQKBEGLNLLSTFYVYKTHFMSNYOHE
ETIYNVLCAPAVNEKFPFRVDGDMVPHISRQLTKYTMADLVALRHFMDGNCOT
LKEILVYTCDDDDYFNKQDWFYFENPDILRVYANIGERVRSLLKTQVQCDAMDA
GIVGLVLDQNLGNWYDFGDEOVAPGCVPIVDSYSLMPLITLALAAESHM
DADLKLPIKMDLLKYDTEERLCIDPRYKWDOTVHNCINCLDRCILKHCANVY
LPSTVPEFTSGFLPKRIPVDCUPFVSTGTHRELGVHVNQDNLNLSRLSKFLLV
YAADPMAHAGNLLDKRTICFSAALTNNVAFQTVKGNFNKDFDFAVSGKPFKE
GSSVELHFFAAGNNAISDIYDYRLPTMCDIRQLFVFEVWPKVITFYDGGGIN
ANQVIVNLDKSGAGFPNKQKARYYDSYEDQDALFAYTKRNIPTITLMLKAYA
ISAKNARTVAGVYSCTMTRNROPHOKLKSIAATRGATVYIGTFKPGFYHNMNLTV
YSDVETPLMGWYPKCDRAMPNMRIMASILARHNTCCNLSHFYRLANCAOVL
SEWVMCGSLYVFGGTSIGSDATAYANVFNICQAVTANNALLSTDGKIKADKYVR
NLQRLUYECLIRNRVDHEFDEFYLRHFSNMILSDDAVVCYNSNAAQGLVASI
KNKAVLYQNNVMSACWETDYLKGPHEFCQSHMLVKQGDYVLYLPDPDSERI
LGAGCFVDDITDGTGLMIERFVSLAIDAYLTKHPNQEYADVFLHYLQIRXKHEL
TGHMLDYSVLTNDTSTRWEPEFYEAMVPTVLOAGCAVCLNSOTSRLCGACIR
RPLCLCKCCYDHIYSTHKLAVLWNPVVCNAPGCDVTDVTLQYLGMSYCKSKHKEPI
SPFLCANGQVFLGYKNTCVGSDNVDFAATCDWTNAGYIILANTCERLKLPAET
LKATETFKLSYGIATVREVLSDRELHLSWEKGRPLNRRNVYFTGYVRVTKSVQOI
GEYTFEKGDDYDVAVYRGTTYKLNVGDYFVLTSHTVPLSAPLVPQEHYVRITGLY
PTLNSIDFSSNVANYQKGMQKYSTLQGPFGTCKSHFAGLALYPSARIVYTAGSH
AAYDALCEKALKYLPIDKCSRIIPARAVECFDKFKVNSTLEOVYCTVNALPETTAD
LVVDELSMATNYDLSVNARLAKHVYVIGDPAQLPAPETLTKGTLEPEYNSVCR
LMKTIIDGDFMLGTRCCPAEIVDTVSALVDNKLKAHKOKSAOCFAMFYKGVITHDVS
SAINRPOIGVVRFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQVDSQSEYDY
VITQTETASHCNVRNFVAITRAKIGILCIMSDDLKQDGLTSLTETPRRNVATIQ
AENVTLFKDCSKITIGLQPTQAPHLSDVIEKTEGLCVDIIPGIMQMTYRRLISNM
GFKMYOVNGVPMFIFREBAIRHVRWVIGFDEGGCHATRDVPLNLOLGFSTGVN
LVAVPGYVDTENNETRVNAXPPGQOKHLLPLMYKGLPNRVRIKIVQMLSDTL
KGLSDRVFVIMAHGFLTSKRYFKIGPBTCLCDKRCATFSTSDSTDCVWNHVSIG
FDYVNFMDIQQWQFTGNLQNDHQHCVGNHVASCDAMTRCLAVHECFVRVYR
DWSVEYIIGDELRINSACKVQHMVYKSLALLKFPVLHDIGNPAIKVQPAQVEM
KGYDAQPCSKAYKIELFYSATHDKFTDGLVFCWNCVDRYIPAKAICVRDFTVL
SNLNRKSCDGSLLVNHKAHPTAFDKSAFTNLKQLPFFVYSDSPCESHCKOVVSDID
YVPLKSACTIRCNLGNACVRHANEYRVIDAYNNMISAGFSYMLYKQFDNTYMLNT
YPLQSLNVAYNVNGHDFHAGEAPVSIINNAVITYKVDIGDVELFENKTTLPNVY
AFELWAKRNTKPIPEIKLNNLQVDAANTVIMDYKREAPAHSTYTCVMTDIAKPP
TESAGSLVTLFDRBEQVQDLFNANRGLVITEGSVKGLTPSKGAQVANGVTILIG
ESVQTFNYFKKVDGIILOQLPETVFTQSRDLEPKRPSOMETDFLEAMLEFQRYKL
EGYAFHEVYDPSHGQGLGLHLMIGAKESQSPLEKDFIPWDSYKNVYFILTDAQ
GSSKCVCSVIDLLDDFVEIJKSQDLSVISKVYKVTIDYAEISFMLACKGHVETFPY
KLQASQAWQPGVAMPNLYKMQRMLEKCDLQNYGENAVIPKGIIMNVAKTYQLCYLIN
TLTLAVPNMRVITHFGAGSKGVAPGTAFLRQMLPTGTLLVDSLDLNDFVSDADSTLIG

DCATVTANKWDLIISDMYDPRTKHVTKENDSKEGFPTYLCGFIKQKLALGGSIAVKI
TEHSNWADYKLMGHFSWMTAFVTNNVASSSEAFILICANYLGPKEQIDGTYTHWANI
FWRNTPIQLSSYSLSLFDMSKFPFLKQVATVMSLKENQINDMIYLSLEKSLIIRENNR
VVSSDILVNN"
249. .13397
/codon_start=1
/product="orf1a polyprotein"
/protein_id="AAR16181.1"
/db_xref="GI:38304882"
/translation="MESLVGVNNEKTHVQLSPVLQVDRVLRGFGDSVEALSEARE
HLKNGTCGLVELEKGVLPOLQOYPVFKSDALSTNHGKVBVLVAEMDGIQVGRSGI
TLGLVPHVGEPIPLAYRNVLLRKNKGAGGHSYGDILKSYDLGDELGTDPIDYEON

Query Match 90.9%; Score 20; DB 14; Length 26333;

Best Local Similarity 75.0%; Pred. No. 0.013;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCGAGGAGCUC 20

Db 760 GTGAACCTCACTCGTGAGCTC 779

RESULT 4

AY463060 29013 bp RNA linear VRL 05-JAN-2004
LOCUS SARS coronavirus ShanghaiQXC2, complete genome.
DEFINITION AY463060
ACCESSION AY463060
VERSION AY463060.1 GI:40457448

KEYWORDS

SOURCE

ORGANISM

SARS coronavirus ShanghaiQXC2
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
1 (bases 1 to 29013)
Yuan, Z., Zhang, X., Hu, Y., Lan, S., Zhou, Z., Wang, H. and Wen, Y.
Analysis of SARS coronavirus genome in Shanghai isolates
Unpublished

REFERENCE

AUTHORS

TITLE

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

location/Qualifiers

organism

mol_type

strain

db_xref

join

notes

codon_start

product

protein_id

translation

length

db_xref

join

notes

codon_start

product

protein_id

translation

length

db_xref

join

notes

VEZEEEDWLDITTEQSEIEPEPPTTPEPVPNQFTYGLKLTDTNVAIKVDUTVREAQOS
NPWYVNAANIHLKHGGVGAALNAKNGAMQSSDDYIKLNGPILKLTDTNVAIKVDUTVREAQOS
AKCLVHVPNLNAGDYLQALNAENFMQSDYILKLPGLSAGIFGAKQPSLOVQCVQT
VRQVYIAENKALYEQVWVYLDNLKPRVEAPQOEPPENTDESTEKSVQKPVDD
KPKEIKACIDEVWTLTVEZFTNLKVLINGLYDQSNMLAGEDMSFLESDAPVM
VGDVYISGSDTCVPIPSKAGQETEMLSRALKAPVDVYIYTPQOCGACQVPEAEK
ALCKKSAFVLPSAPENAKBEETLCTSVNLSREMLAHAEERAKMLPCMDVRAIMATI
QRKYKGVLOQSIGVDFRFFVYTSKEPVASIIITKLSNLEPLVTMIPGVTHGPNLE
EAKRMESKLAPAVSVYSDPVTYVINGVLTSSKTSIEHFVTSVLAGSTRWVSYG
QRELGEVPLKAGDVYVYHSDPEVYHLDGVSLEDKLKSLEUREVYIKVTVVYD
YTNHLQLOQMSVYTCQOPEPTVLDGAYTKIKPHNVNCEGFTFVLPJSDTDLSEAFEP
YTNHLDESFGYMSALNTKKWPOUGVTSIKWANDNCCVLSVALLAQOLEKVENP
APALQEAYRARGADANFICALIAYSNKTVGELGDVRETHWLOHANLESASKVLN
VCHQCOQTLITVEAYNMGTSYDNLKTGYSIPCVCBGRDATOYLQOSESFVMM
SAPFAGKLOQGFPCANEYITQGHYHTKATSIYPCVCBGRDATOYLQOSESFVMM
VYKETSYYTIIKPVSVKUDGVYIETPELQDGYKKONDAVENTSIITKPNLSLFT
LGKTAITHGAIANINPMVKILAYKPELQOAIITNSCARLQAVFNFNYPYFT
LLLOLCFTFKTSNSTRASLPITLANSKVCYKAKLDAGNVYKSPKFLCPIAMM
LLLSILCGLSILCTVAFGLPITLANSKVCYKAKLDAGNVYKSPKFLCPIAMM
CLSGLOSDOSYPALEITQVTSYKLDLTILGLAAEWLYAMLPKTFYLLGLSALIM
AFQGFASHISNSLWMLPIISYVQWAPSVAMVYIIFASPYIYKSYHIHMDGCTS
STCMYCKNHAIRVECTIYNGMKSFVYVANGRFCKTHNCLNCDITCTGTSFT
ISDEVARNLSLOKFRINPTDSSYVSDVAGNALHYFDPAQKTYTHERPLSHFV
NLNDRANNTKSGSLNPIYVDFCKSGDESASKASYVQLOMCOILLQDQVSDV
GDSYTESWKFADVYDTSATSPMEKILKALVATYKVALGVDQISVFSAR
QGVYDVTQDOVIELCKLSHSDILEVTEGSCYNPMLTYKNVENMTPRDLSACIDNCA
RHINAQVASHNVSLIMVYKYSLSBQLKQIRSAAKKNIPELCTATQVQVNI
TYTKLSGGKIVSTCFKMLKATLVCALVYIYVMPHTVTHDGTNTEIIGYKAI
QDGVYDITSDCFARGHAGDFAMFSQSGYKNDKSCPVVAIIITREIGYVGLP
GYLRAINGDLPLRFVSAGVNTCYTPSKSIIQYDTSATSAALAEETFKDAMK
PVVYQVNTLLEGSIIYSVELPDPTRVYLMKSLIIOFNPYTEGVSVRVVTDFAYCRH
GCTERSBGIKSGELSGSGVFNIDVDSYCFVMMHLETPGVGHATDLEGKFGPVDD
ASVQAGYIALLVTCRAVYEMFERAAGEYHVAANALLFMSFTILCLAVPASELP
GVYSVFLYITFYFNDVSELAHLQWFAMFSPVPHPIYAVPCLSKHCHWFNNY
LKRKVMFGVYTSFEAACTFLNKEMYLKRSLETPLTPQNYRLXKMLTKYFS
GLVDTYSREAAACHAALNDFSNGSDVLYQPOPTSITSALVQSGFKWAPSGK
EGCVQVTCETTLMLDLDYVTSYPRHYICTAEDMLNPNEDLIRKSHNSLVAQ
NQGLVYTCGTMQNLRLKLDYVTSNPKYKEVRITQOCTESVLACNGSPSVYQAC
NPNRHTIKSGELSGSGVFNIDVDSYCFVMMHLETPGVGHATDLEGKFGPVDD
RQNTAQAAGTPTTITANVLAALVAAYDGSFCVMMHLETPGVGHATDLEGKFGPVDD
HVLLGLPSAQGLIAMDCAAKELLONGNGRTIGSTLEDBEFTPPVQRCQSGV
TQGRKFIYKGTWHMLLTITLSLIVQSTWQVPMWMLTLEADTSLSGYRLKDAC
AMLLVHKHAFCLCLPLSATVYVNMVPMWMLTLEADTSLSGYRLKDAC
VMTASALVITLMEFATVYDAARVYTMVNIITLYKMYVGNALQAIISWALVISA
TSNYSGVITIMEFARAIYVFCVEYPLFTITQCLMYVAGQCCSPFGFLC
LLNRYPSLMDVSVLVSQFVYKNSGLYIKGLNGLNMGVLSLQATVLOAGNATEVPA
LVQSGSDVKVTSVLSVQFVYKNSGLYIKGLNGLNMGVLSLQATVLOAGNATEVPA
TSVLLSHQAGVDNLNGLCEVILNRAILOALASESSLPSAAYATAQBYEQAVANGD
SEVVLKRLKJDNALNINARDQGVPLIIPDQAAMQVYDGYKNTCDGT
TMFLTMLKJDNALNINARDQGVPLIIPDQAAMQVYDGYKNTCDGT
VYASALWETQQVDAKCIQVSEINMDSNPALVITLARSAAKLNNESSLP
VALRMSCAAGTQTDONALAYNNSKGRFVALLSDHQLKWARPFGSDGTGT
YITELBPCHPFTVDTQPKMVKLYIKGLNGLNMGVLSLQATVLOAGNATEVPA
NSTVLPSAFVDPKAYKDYLAGSOPQVNCVMLCTHGTQATVTPRANNDGDS
FGASCLCYCRCHIDHPKGFCDLKGKIVQIPTTCANDPVGFTLRNCTVCMWKG
VGCSDOLREPLMSADASTFKRVGVSAAKLTCGTSITVYVAREYNEKVG
FAKPLKTRCNQFQDEBGNLLYVFKWRTMSNYCHEETIYNLIDCPAVALHFF
KFRVDGMWPHILSQBLTKYTWADIVYALRHDEGNCDLGLIYVNCDDDYFNKK
VDFVDEBNPDILSRALNGLBERVSLQTLVQFCADMRADGIVGLVTDONLGNWYD
EGRFQVQAPCGVPIYDSYVSLMLPILTLTRALAAESHMDADLAKPLKMDLKYDPT
FEEGLIFDRYFKYVWDTQYHPCNIDBCLILCHANFVLSFTVPTSPGTSGLPVKRI
VDGVPIFYVGTGTHRELQVHNDVNLDSRILCHANFVLSFTVPTSPGTSGLPVKRI
TCTFESVALNNVAFQVTPKFNKDFDFAVKGSGFKNGSSVELKHFFFAQDNKA
SDYDYRNLNPTMCDITROLLFVYDVKDFCGGCLNPNYVNLNLSGAPPNK
WKGARYLSDMSYEQDQALFATYKRNVTIYPTQMLKTAISAKNRATAGVSTCTM
TNRQFOHLKLSAAITRGATVYTSKFGVGHNMMLKTVYSDVETHLMGWYDKPCR
AMPNMLIEMASIVLAKHNTCNLSHREYRLANEAQOVLSEWVWCGSLYKPGSTH
GDATAYANSVFNICQVATNVALLSTDNKILADKYVNIQHLRECLYRNDVDSH
FVDEYIYLRKHSMLSDDAVCVNSYNSAAQVLVSKNFKAVYQNNVFWSEAK
CWETDITKGPHEFCSQHTMLVKQDGYVYLPDPDSILAGAGVDDIKVGTGML
ERFVSILADIPYTCFTHQNSYADVHLYQYIRKLDELGTGHTDMLVMTLNDNSTR
YWEPEFYEAUTHTVQVAGVCLCNSYLSRCCKSHKRPISPELCAONVGLYKNTCV
ILVSNPVCAVPGCDVTQVLOILAGMSYSCCKSHKRPISPELCAONVGLYKNTCV

```

GSDNVDPFNALATCDWTNAGDYILANTCTERBLKLPAAETLKAETEETFKLSVGIATVRE
VLSREKLHSEWEKGVKPTNLRNFTVGIYRVTKNSKQVIGEYTFEKDGIYKLSVGIATVRE
TTYXLANVDGVYKLTSHTPMLPSALVLPQVHYVRI TGLYPTLAINTEFSFNNAVYQVY
GMQKSYTLQGPPTGKSHFAI GLALYYPESARI VYTACSHAADALCEKALKYLP IDRC
SRI IIPASARVBCFDKFKVNSITLEQYVFTCNALPETTADI VWFDEISMATYDL SVNNY
ARLAKHYVYIGDPAOLPAPELTLLTKGTLEPFSVNSVRLMKTIGPMFLGTGCRCPA
EIVDTVSLVYDNLKAHKESACQFMFKYKGVITHDVSSA INRPLOGVVREFLTRNP
AWRKAVIPSPYNSQNAVASKLGLPTQTVUSSQSESDYVI FTQTETAHSCHNVRNP
VAITRAKILICILNCRDLVDKQFTSL EIPRNVATLOQAENVFTFKPQCSKILITGLH
PTQAPTHLSVIDIKFTBGLCDVLP GIPKDMTYRRLISMMGFKNYQVNGVPMFI TRE
BAISHVRPAAGFDGVCHATRDVNLPLQLGFGSTVNLVAVPTGVYVDTENNTEFTL
VNAKPPEDGDFKHLIPLMYKGLPWNVVR I KIVOMLSDTLKGLSDRVVFLVMAHGEFLT
SMKYFPIGPBTRCTCLCBKATCTSGSDTVA CNHNSVGFYVYVNPEMI DVQOMQFTG
NLOSNDHOHQCHNSAHVASCDA LMTFCLAHVECFKRVDMSEVEYPI LGDELAVNSAC
RKVOHMYKKSALLADKPEFVLHDIGNPKA I KCVQAEVEMFKVEDAQPCSDKAYKIEELF
YSYATHHDKFTDGVCLFWNCNVORI PANAI VCHRGDTRVLNMLPGDGGSL VNRHGA
FHTPAFKSAFTNLQJLPFFYYSPECSECHQVQSDI DYVPLKSASTICTRNLGGLAV
CRHIANEYQVYDLAYNNMI SAGFSLYIKQFTYVNLNMTFTRLQSLSENAYVNVNKGH
PDGHAGBAPY I INNAVYKVDGIDVEI FESKTYL PVNVA FELWAKMI KPVEPI KIL
NVLGDVAANTV I WEGSVKLEAPHSVSTIGCTMTDIA AKKTESACSLYPKVDGQVGO
VLD FPNARNGVILITDEKSAKGLPKSGQAQASVNGVTL I GESVKTQFNYKPKDGI I LOQ
LPEYFTQSRDLEDKPRSQMETHD FLELANDBEFTQRYKLEGYAEHLIYGDGFSHGQLG
GLHMLIGLAKRSQVSKPLKED FPMDSCKDNXYETFPYKLOQSAQWQVAMPNLPYK
I IKQSDLSVTSKVYVTVLIDYAEIS TMLDCKDGHVETFPYKLOQSAQWQVAMPNLPYK
MORMLSEKDLQNGENAVI PKGIMNNVAKYITQLCQYVNLTLTAVPYNNRVIHFHGA
DKGVAGPTAVLRQWLPTGLLVDSDLNDVFSDDADSTLIGDCATVHTKANWDL I I SDMY
DPRTKVTKENDSGEFTFYLCGFI KOKLALGGS I AVKI TEHSNADPLYKLMGHFSW
TAFVTVNASSSEAFGLIYANVGLCKPEQIDGYTHANY I FWRNTDPLQLSSYSILFDM
KPPKLARGTAVMSLKENQINDMI YSLLEKGRLLIRENNRVVUSSDILVNN"
20834..24601
/codon_start=1
/product="spike"
/protein_id="AAR86775.1"
/db_xref="GI:40457450"
/translation="MFTFLFLRLTSGSLDRCTTFDDVQAPNVTQHTSSMRGVVYVD
EIFRSDLYLTQDLFLFFSYNVGTGHTINHTGNPVI PFKDGIYFAATEKSNVVRGNV
PEYTSNDFSUDSKSGNFKHLREFVFNKXDGFLYVYKQPI IDVVQRDLPSGNTLKP
I FKLPLQINTFNRALITFAFTSPAQDITFGSAAAYFVYGLRPTFMILKIDGENTATDAV
DCSQNPALAEKLSFKSAFEI DKGIYQTSNFRVPSGVDPNPNTI PKCFGEVFNATKP
PSVYAWERKKI SNCVADSVLNDPDMFCVCLAWNTGNI DATSTGNVNYKYR YLRHGLRP
DVROI APQGTGVIADNYKLPDOPMFCVCLAWNTGNI DATSTGNVNYKYR YLRHGLRP
FERDI SNVPSRPGCKPCTPALANQVPLNDYGFYTTTG IGYQPVVVVLSFELNAPA
TCQPKSLTDLILKNQCVNFGNLGTGTVLTPSKRPPQPPQPPQPPQPPQPPQPPQPPQPPQ
TSEILDSPACFGSVNITPGTNASSVAELYQDVNCTNVSSAAIHADQLTPAKRI VSTI
GNVYFQOAGCLIGAEHVDSITPEIGAGI CASYHTVSLSLSTSQKS I VAYTMSLG
ADSSYATSNNTIATPTNFIS IITEVNPVSMAKT SVDCNNY I CGDSTECANLLQYGS
PCTOLNALSIGIAEADQNRTRFEAVQVQMKQPTLKYFGGFNFSQLLPDLKPTKRS
PTIEDLLNGLTADQDPMKQVCEGICLDNARDLI CAQFNGLTVLPDLLTDMLAAVT
AALYSGTATAGTFCGAALQIIPAMOMAYRFGNGIYGVNTQVLYENKOI I ANQFNKA
QIQESLTTTALGKLQDVNVNQNALNITLVKQLSNFGLASVVLNDI I LSRLDKVEAE
VQIDRLITGRLSQIQTIVTQOLRAABIRASANTAKMSECVGLQSKRVDP CKGHTH
LMSPPQAPPHGVFELHYVPSYQPERNFTTAPAI CHEKGYAPREGVFNFGTSWFTIQ
RNFFSPQ I ITDNTFVSGNCDVWIGI I NNTVYDLPQPELDSFKGELDKYKPNKSTPDV
DLGDI SGINASVNI QCEI DRLNEAVANL NESLI DDEIQLGKYEQY I KPWVYWLGIATA
GL I LVMYTL I LCVMTSCCLKGCACSCGSCDKFEDDSEPVLYKKVHYVT"

```

CDS

CDS

```
24610..25434
/codon_start=1
/product="unknown"
/protein_id="AAR86776.1"
/db_xref="GI:40457451"
/translation="MDLPMRFPTTGSITAPQKIDNAPASTAHATAT1PQASLSPGG
WLIVGVAFVAFQSAQKLIALLKRWQALYKGFQFCNLLLLFTVIYSHLLLAAGME
AQFLVLLALVQLQCINACRIIMRWLCWKCKKNPLLYDANFVCMWHTHNYDICYIPY
NIVTGTIGVTEGDSIGTPELKEDYQIGGYSEDRHSQVKVXVVGHTFEVYVQLESTQ
ITNTDGTIENATFTFENLKVDPNNVQIHTIDGSSGVANPAMPDIYDEPTTTTSVPL"
25031..25495
/codon_start=1
/product="unknown"
/protein_id="AAR86777.1"
/db_xref="GI:40457452"
/translation="MPMTTIFAGTHITMTTVYTHITVTSQIQLSLLKVTAFOHNSKKT
KLWILRIGTOLVTKMSLYMAISPKEFTLQTLVLKMLHSSSI2TSLKATHRM
```

CDS

CDS	<p> CKYQTSTALQELLQQIMQFMSSRRRLACLCKKKKYSTNLCTHSFRKKQVR" 25459. .25689 /codon_start=1 /product="E" /protein_id="AAR86778.1" </p>									
Query Match	90.9%; Score 20; DB 14; Length 29013;									
Best Local Similarity	75.0%; Pred. No. 0.013;									
Matches	15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;									
QY	1 GUGAACUCACUCUGGAGCUC 20									
DB	697 GTGAACCTACTCGTGAGCTC 716									
LOCUS	AY394999 29350 bp RNA linear VRL 29-JAN-2004									
DEFINITION	SARS coronavirus LC2, complete genome.									
ACCESSION	AY394999									
VERSION	AY394999.1 GI:37624342									
KEYWORDS	SARS coronavirus LC2									
SOURCE	SARS coronavirus LC2									
ORGANISM	SARS coronavirus LC2									
REFERENCE	Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales; Coronaviridae; Coronavirus.									
AUTHORS	1 (bases 1 to 29350)									
CONSTRM	The SARS epidemiology consortium of Guangdong									
TITLE	From independent foci of epidemic outbreak to large genomic alteration in late phase viruses: evolution of the SARS-coronavirus									
JOURNAL	Unpublished									
REFERENCE	2 (bases 1 to 29350)									
AUTHORS	The SARS epidemiology consortium of Guangdong									
CONSTRM	Direct Submission									
TITLE	Submitted (19-SEP-2003) Guangdong, China									
JOURNAL	Location/Qualifiers									
FEATURES	1. .29350									
source	/organism="SARS coronavirus LC2"									
	/mol_type="genomic RNA"									
	/isolate="LC2"									
	/db_xref="taxon:249084"									
ORIGIN										
Query Match	90.9%; Score 20; DB 14; Length 29350;									
Best Local Similarity	75.0%; Pred. No. 0.013;									
Matches	15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;									
QY	1 GUGAACUCACUCUGGAGCUC 20									
DB	761 GTGAACCTACTCGTGAGCTC 780									
LOCUS	AY395000 29350 bp RNA linear VRL 29-JAN-2004									
DEFINITION	SARS coronavirus LC3, complete genome.									

```

TITLE      From independent foci of epidemic outbreak to large genomic
JOURNAL    alteration in late phase viruses: evolution of the SARS-coronavirus
REFERENCE  2 (bases 1 to 29350)
AUTHORS
CONSTRM
TITLE      The SARS epidemiology consortium of Guangdong
JOURNAL    Direct Submission
JOURNAL    Submitted (19-SEP-2003) Guangdong, China
FEATURES
source
ORIGIN
Query Match      90.9%; Score 20; DB 14; Length 29350;
Best Local Similarity 75.0%; Pred. No. 0.013;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1 GUGAACUCACUCGUGAGCUC 20
       |:||||:||||:||||:|
Db      761 GTGAACCTCACTCGTGAGCTC 780

RESULT 9
AY394977
LOCUS      AY394977      29433 bp      RNA      linear      VRL 29-JAN-2004
DEFINITION SARS coronavirus GZ-A, partial genome.
ACCESSION  AY394977
VERSION     AY394977.1 GI:37624320
KEYWORDS
SOURCE
ORGANISM   SARS coronavirus GZ-A
           SARS coronavirus GZ-A
           Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
           Coronaviridae; Coronavirus.
REFERENCE  1 (bases 1 to 29433)
AUTHORS
CONSTRM
TITLE      The SARS epidemiology consortium of Guangdong
JOURNAL    From independent foci of epidemic outbreak to large genomic
JOURNAL    alteration in late phase viruses: evolution of the SARS-coronavirus
REFERENCE  2 (bases 1 to 29433)
AUTHORS
CONSTRM
TITLE      The SARS epidemiology consortium of Guangdong
JOURNAL    Direct Submission
JOURNAL    Submitted (19-SEP-2003) Guangdong, China
FEATURES
source
ORIGIN
Query Match      90.9%; Score 20; DB 14; Length 29433;
Best Local Similarity 75.0%; Pred. No. 0.013;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1 GUGAACUCACUCGUGAGCUC 20
       |:||||:||||:||||:|
Db      725 GTGAACCTCACTCGTGAGCTC 744

RESULT 10
AY394985
LOCUS      AY394985      29530 bp      RNA      linear      VRL 29-JAN-2004
DEFINITION SARS coronavirus HSZ-Bb, complete genome.
ACCESSION  AY394985
VERSION     AY394985.1 GI:37624328
KEYWORDS
SOURCE
ORGANISM   SARS coronavirus HSZ-Bb
           SARS coronavirus HSZ-Bb

```

```

Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
1 (bases 1 to 29530)
The SARS epidemiology consortium of Guangdong
From independent foci of epidemic outbreak to large genomic
alteration in late phase viruses: evolution of the SARS-coronavirus
Unpublished
2 (bases 1 to 29530)
The SARS epidemiology consortium of Guangdong
Direct Submission
Submitted (19-SEP-2003) Guangdong, China
Location/Qualifiers
1. .29530
/organism="SARS coronavirus HSZ-Bb"
/mol_type="genomic RNA"
/isolate="HSZ-Bb"
/db_xref="taxon:249070"
source
ORIGIN
Query Match      90.9%; Score 20; DB 14; Length 29530;
Best Local Similarity 75.0%; Pred. No. 0.013;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY      1 GUGAACUCACUCGUGAGCUC 20
       |:||||:||||:||||:|
Db      526 GTGAACCTCACTCGTGAGCTC 545

RESULT 11
AY338174
LOCUS      AY338174      29573 bp      RNA      linear      VRL 28-JUL-2003
DEFINITION SARS coronavirus Taiwan TC1, complete genome.
ACCESSION  AY338174
VERSION     AY338174.1 GI:32493129
KEYWORDS
SOURCE
ORGANISM   SARS coronavirus Taiwan TC1
           SARS coronavirus Taiwan TC1
           Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
           Coronaviridae; Coronavirus.
REFERENCE  1 (bases 1 to 29573)
AUTHORS      Chang,J.-G.C., Lin,T.-H., Chen,C.-M., Lin,C.-S., Chan,W.-L. and
              Shih,M.-C.
TITLE      SARS coronavirus TC1, clinical specimen
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 29573)
AUTHORS      Chang,J.-G.C., Lin,T.-H., Chen,C.-M., Lin,C.-S., Chan,W.-L. and
              Shih,M.-C.
TITLE      Direct Submission
JOURNAL    Submitted (08-JUL-2003) Department of Molecular Medicine, China
              Medical University Hospital, 2, Yuh Der Road, Taichung, Taichung
              404, Taiwan
REFERENCE  3 (bases 1 to 29573)
AUTHORS      Chang,J.-G.C., Lin,T.-H., Chen,C.-M., Lin,C.-S., Chan,W.-L. and
              Shih,M.-C.
TITLE      Direct Submission
JOURNAL    Submitted (28-JUL-2003) Department of Molecular Medicine, China
              Medical University Hospital, 2, Yuh Der Road, Taichung, Taichung
              404, Taiwan
REMARK     Amino acid sequence updated by submitter
FEATURES
source
1. .29573
/organism="SARS coronavirus Taiwan TC1"
/mol_type="genomic RNA"
/isolate="TC1"
/db_xref="taxon:237639"
/note="clinical specimen"
join(196..13329,13329..21416)
/codon_start=1
/product="nonstructural polyprotein pplab"
/protein_id="AAQ01594.1"
/db_xref="GI:33285967"
CDS

```

/translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRCFGDSVEEALSEARE
HLKNGTGLVELEKGLPQLEQPIYFKRSDALSTNGHKKVVELVLAEMDGIQYGRSGI
TLGVLPVHVEGTPIAYNRVLLRKNKGAGCHSYDILKSYDLGDELGTDPIDEDYEN
WNTKHSGALRETLRELNGAVTRYVDNNGFCGPDGYPLODCIKDFLARAGSMCTLSQQ
LTYIDESKGVVCCRDEHHEHIAVTFERSDKSYEHOTPEIKSAKFFPKCEKCPKFPV
LNSKVYIOPQREKKTGEGPMGRTRSYVVPASPOECNNMHLSTLMKCNHDEVSMTQC
DLFKATCEHCCTENLVEGPTTCGYLPTNAVWKMPACADPELGPBHSVDYHNHNS
IETLRKGRTRCFGCYFATYGCINRATYVPRASADIGSGHTGIDVNDVETNDEL
LEILSRKRVNINIGDPLHNEEVAIILASFSASAFIDITIKSLDYKSFYTIYESCGN
KYSYTVESGTEVTEFAEVAEAVVTKLPQVSDLLTNMGIDLDEWSEYLFYFADAGE
ENFMSKSYSPPEDEEDDAECBEEIDBTECHBYGDEHVDQGLPLEFAGSAETVR
VEEEREDMLDDTTEQSEIRPEPEPTPEEPNOFTGYLKTNDVAIKCVDIVKEAOSA
NPWIVNAAINHLKHGGGAGALNAKTAGMQRKESDDYIKLNGPLTVGGSCLLSGHL
AKKLVHVPNPNLHAGGDIQLLKAAYENFNODILLAPLLSAGIPGAKPLQSLQCVOT
VRTQYIAVNDKALYEQVMDYLDNLKPRVEAPKQEPNPTEDSKTEEKSVMQKPVV
KPKIKACIDEVTEFTLTKLNLKLLFADINGKLYHDSQMLRGESEFLEKDAFYM
VGDVITSGDITCVVIPSCKAGGTTEMLSRALKKVPVDEYITVPGQCGAGYLEEAAT
ALKKKSAFYVLPSEANAKBEEILGTVSMNREMLAHAEKTRKMLPIMCDBRAIMATI
ORKYKGIQEGIVDYGRRFFFTSKSPVASIITKLSNLEPLVTMPIGIVTHGFNLE
EARKCMSLKAPAVSVSSPDVAITYNGYLTSSKTSSEHFVETSUAGSYRMSYSG
ORTLEGYFLKRGDKIYHTLESPEFHLDGEVLSLKLKJLSLSREKVIKYFTVTD
NTNLLQIDVMSMTYGOQFGPTLYDGDADVTKIPHNHEGKTFVLPDSDTLRSEAPE
YHTLDSFLGRYMSALHNTKWKFPQVGGITSIKWANDNYLSVSLALQOLEKFN
APALQEAAYRARGADAANFICALLAYSNKTVGELGDVRETMTHLLOHANLESKVLN
VVCHEQKQTTLLTGVAZVMGTSLDNLKTVGSIIPCVQGRDQATLQVQESFVMV
SAPPAEYKQOQTEFLCANEYTGNYQCCHYTHIYAKETLYRIDGAHLLTKQESKGPVD
FYFKETSYTITKEPVSJLDGVTYTEIEPKLDGYSYKKNDAVYTEQPIIDLVPQLPNA
SFNFKLTCSNTKPADLNOAGMTGFKPASRELSVTFPPDNGDVVAIDYRHYSAPKKA
GAKLLHPIVMIHINOITKTTEKPNWCLRWSTKPVDSNSFEVLAVDTQGMNLI
ACESQPTSEEVNPTIQKEVLECDVKTTEVWGNVLKPSDEGVKVTQELGHEDLMA
AYENTSITIKWPELSLALGLKTIATHGAIANSVSPKSLATVFPFLQQAALTSN
CAKRLAGFNNMYPTVTLFOLCTFKTSNRSIRASLPTTIKANSVSGYKACILDA
GINYKSPKSPKTIAMWLLLSICLSLICTVAAPGVLSSNFGAPSGYNGVRELYL
NSNNTVTFDCEGSPFCSICLSGLSDSPALPETIQTIVTSYKLLDITGLAEAWLI
AYMLTFYFLYLGSLAMOVFGYFASHFISNMLWFIIISIQMAPVSAVMRYIIF
ASPHYIKSYVHINDGCTSCMCKYENRATRECTIUNGKRSFYVYANGRGFC
KTHWNGCNDCTGCTFISDEVARDLSQFKRPINPTQDQSTIVDSVAKNGALHL
YFDKAGQTRERHPLSHFNLNLRNANTKSLSPINIVFDGSKDSKESKASVYV
SQLMCQILLDQALSDVDGDSVEVSDFDAYVDLTFATSPFMEKMLKALVATAHSE
LAKVADLGVLSLTPVSAARQGVTDVTDKDVIBCLKSHSDLEVETGDCNNFMFLTY
NKVENMTPRDLGACIDCNARHINAQAKSHNSVLIWNKDYMSISEOLRKOIRSAKK
NNIPRLTCATTRQVNVITTKISLKGKIVSTGFKMLKATLLCVLAALVYCIYMPV
HTLSIHGYNEIIGYKAIQDGVTRDIISTDDCPANKHAGFDWFSORGSGYKNDKSC
PVAAIITREIGTIVPGLPGLTRAINGDFLHPLPRVFSAGNICTYTPSKLIYESDFA
TSACVLAECITFKDAMGKVPYCYDNTLLGSGISYSELAPDTRYVLMDCGSIIOFPNT
YLEGSVRVTTFDABYCRHGTCESEVGI CLSTSGRWLVNNEHVALSGVFCGVDMAN
LIANIPTLPQPCALDVSASVAGGIIAILVTCAAYYFKERPRVGEYNHVVAAANL
LFLMSFTILCIVPAYSFLPGYSVYFYLIFPYFTNDVSVFLAHQWFAWFSPIVFPWIT
AIYVFCISLKHGFFNNYLRKRVNMGVNTFTFEAAALCTFLINKEMYLKRSSETLL
PLQYNYRYLALYNYKTYFSGALDITSYREAAACHLARDNFNSGDVILYQPPQTSI
TSYDLQIRKSNHSPLOAGNVOLRIVIGHSMONCLLRKLVDTNPNKPKYKFRVIOPOQ
TESVLA CNGSPSGVQCMRPNHTIKGSPFNGSCSGVGNIDYCVSCFYCMHMEPL
TGVIAGDTLSEKFPYGPVDRQTAAGDITITILNLAWLVAAYVINDGRFLNRFTTT
LNDPFLVAMKRYNEPLDQHDVILGLPSAQTGIAVLDMAKALLKELLQNGWNGTILGS
TILDEBFTPTFDVAVRQCSGVYFQGGPKKIYKGTTHMULLTFLTSLILVQTSLSLFPF
VYENAFPTFTLGINAIIAACAMLAVKHKAFLCLFLPLSLATVAYFNNVMPASWVRI
MTLEADTSLSGYRLKDCVYASALVLLIIMTARTYVDDAARRVYTMVIVLYYKV
MYGNALDOALSMWALVSVTSNGVYTIMELARAVFVCEVYPLFLTGTNLTQCI
MLVYCFYCCYCCYFGLPCLLNRYFLRLTGLVYDLVLTQBFRTMNSGGLLPKPKSDI
FKMLKULGIGKPKCIKAVIQSKSDVKCTSIVLLSVLOQLRVESSESKWAQCQLH
NDIILAKDPTTEAFKRYVSLLSVLSMQGADVINRLCEMDLNRTLOAIESFSSLP
YAAVATAQEAQAVANGDSEVLKCLSNVAKSFEDRDAQMRKLEKMAQDIAK
MYKQARDEKAVTYSAMOTMLFTMLRLKLDNALNNTINARDQCPVLNIIPLTAAQ
LMVVPDYGTYNKDCGNTFYASALHEIQVVDADSKIYQLSEINNDNSPNLAWPLI
VTALRANSAVKLQNNELSPVALROMSCAAGTITQATCDNALATYNNKSGRFLALL
SDHQDLKWARPKPSDGTGTIYTELEPPCRFVTDPKGPVKYKLYFIYIKGLNNLRNGWL

CDS

/codon_start=1
/protein_id="AAQ01595.1"
/db_xref="GI:33285968"
/translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRCFGDSVEEALSEARE
HLKNGTGLVELEKGLPQLEQPIYFKRSDALSTNGHKKVVELVLAEMDGIQYGRSGI
TLGVLPVHVEGTPIAYNRVLLRKNKGAGCHSYDILKSYDLGDELGTDPIDEDYEN
WNTKHSGALRETLRELNGAVTRYVDNNGFCGPDGYPLODCIKDFLARAGSMCTLSQQ
LTYIDESKGVVCCRDEHHEHIAVTFERSDKSYEHOTPEIKSAKFFPKCEKCPKFPV
LNSKVYIOPQREKKTGEGPMGRTRSYVVPASPOECNNMHLSTLMKCNHDEVSMTQC
DLFKATCEHCCTENLVEGPTTCGYLPTNAVWKMPACADPELGPBHSVDYHNHNS
IETLRKGRTRCFGCYFATYGCINRATYVPRASADIGSGHTGIDVNDVETNDEL
LEILSRKRVNINIGDPLHNEEVAIILASFSASAFIDITIKSLDYKSFYTIYESCGN
KYSYTVESGTEVTEFAEVAEAVVTKLPQVSDLLTNMGIDLDEWSEYLFYFADAGE
ENFMSKSYSPPEDEEDDAECBEEIDBTECHBYGDEHVDQGLPLEFAGSAETVR
VEEEREDMLDDTTEQSEIRPEPEPTPEEPNOFTGYLKTNDVAIKCVDIVKEAOSA
196..13344
/protein_id="nonstructural polyprotein ppla"
/db_xref="GI:33285968"

NPWVIVNAANIHLKHGGVAGALNKATNGAMQKESDDYIKLANGPLTVGGSCLLSGHNL
AKKLVAVGNLNAAGSDIOLLKAAVFNFSQDILLAPLISAGI FGAKPLQSQVCQVOT
VRTOVYVNDLKVQVMDYLDNLKPRVAPKQEPENPTNDSKTEEXSVQKPDV
KPKAKACIDVETTLTETKFLTNKLLLFADINGKLKHDSONMLRGDMSFLEKDAAPYM
VGDVITSGDITCVIIPSKKAGGTETMLSRALKKVPVDEYITTPGGCGAGYTLSEAKT
ALKKCSAFYVLSEAPNAKEBILGTVSNLREMLAHABETRLKMPICMDVRAIMATI
ORRYKGIKTQEGIVDGVRFPPFTYTSKPEVASIITKLSNLEPLVTMPICMDVRAIMATI
EAKRMSRLKAPAVSVSPDATTYNGYLTSSKTSSEHFEVTVSLAGSYRDSYSG
ORTELGVFLKRGDKIVHTLESPEFHLDGEVLSLDKLSLSLSEVKTIVKPTTVD
NTNLHTQVDMSTMKYQQGFTYLDGADVTKIKPHVNHGKTFVFLPSDDTLRSEAFE
YHTLDESFLGRYMSALNHTKWKFPQVGLTSIKWADNNCYLSVLLALQQLSEAFN
APALQAYYRARGDAANFCALILAYSNKTVELGVDRETMTLHQLHANLESKRVLN
VVKHCGQKTTTLTGVEAVMYMGTLSDYDLNKTGVSIPCVCGRDATQYLVQOESSFVMW

Query Match 90.98; Score 20; DB 14; Length 29573;
Best Local Similarity 75.08; Pred. No. 0.013;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCGUGAGCUC 20
Db 707 GTGACTCACTCTGTAGCTC 726
|:||||:||||:||||:|
|:||||:||||:||||:|

RESULT 12
AY338175 29573 bp RNA linear VRL 28-JUL-2003
LOCUS SARS coronavirus Taiwan TC2, complete genome.
DEFINITION SARS coronavirus Taiwan TC2, complete genome.
ACCESSION AY338175
VERSION AY338175.1 GI:32493130
KEYWORDS
SOURCE SARS coronavirus Taiwan TC2

ORGANISM SARS coronavirus Taiwan TC2
SARS coronavirus Taiwan TC2
Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.

REFERENCE 1 (bases 1 to 29573)
Chang, J., Lin, T., Chen, C., Lin, T., Chan, W. and Shih, M.
TITLE SARS coronavirus TC2, clinical specimen
JOURNAL Unpublished
AUTHORS Chang, J., Lin, T., Chen, C., Lin, T., Chan, W. and Shih, M.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2003) Department of Molecular Medicine, China Medical University Hospital, 2, Yuh Der Road, Taichung, Taichung 404, Taiwan

REFERENCE 3 (bases 1 to 29573)
Chang, J., Lin, T., Chen, C., Lin, T., Chan, W. and Shih, M.
TITLE Direct Submission
JOURNAL Submitted (28-JUL-2003) Department of Molecular Medicine, China Medical University Hospital, 2, Yuh Der Road, Taichung, Taichung 404, Taiwan

REMARK Amino acid sequence updated by submitter
FEATURES Location/Qualifiers
1. 29573
/organism="SARS coronavirus Taiwan TC2"
/mol type="genomic RNA"
/isolate="TC2"
/db xref="taxon:237640"
/note="clinical specimen"
join(196..13329,13329..21416)
/codon start=1
/product="nonstructural polyprotein ppiab"
/protein_id="AA001606.1"
/db xref="GI:33285979"
/translation="MESLVGVNKEKTHVQLSLPVLQVRDLVVRGFGDSVEEALSEARE
HLKNGTCGLVEKSGVLPQLEQYVFIKRSDALSTNHGHVKVVELVAMGDIQGRSGI
TLGVVPHYGETPIAYVNLRRKNGKAGHSYGLDKSYDLGDLBGLDPTDEYQON
WNTKSGSLRELRETLNGAVTRYVDNNPFCGPDGYPLDICKDFLAGRKMCTLSEQ
LDVIESKRGVYCCRDHEHEIAWTFERSDKSYEHOTPEIKSAKFTFKGCKFPVP
LNSKVKIOPRVESKKTTEGPMGRISVYPVAPSQECNWHLSLTMKNCDEVSQWTC
DFLCAECGCTENLIEGPTTCGYLPTNAVVMPCAPCDPBGPEHSHVADYHNHNS
IETLRKRGKRCFCGGFYAYCYNKAIRYVPRASADIGSGHTIGTGNVEDSL
LETLSRVENINTVGFHLNEEVAIIAGFSASTSAPIDTILSKDYSKFTIVESCGN
YKVTGKPKVGAWNIQOQRSVLTPLCGFPSSQAAGVIRSIPARTLDAANHSIPDLQRAA

CDS
join(196..13329,13329..21416)
/codon start=1
/product="nonstructural polyprotein ppiab"
/protein_id="AA001606.1"
/db xref="GI:33285979"
/translation="MESLVGVNKEKTHVQLSLPVLQVRDLVVRGFGDSVEEALSEARE
HLKNGTCGLVEKSGVLPQLEQYVFIKRSDALSTNHGHVKVVELVAMGDIQGRSGI
TLGVVPHYGETPIAYVNLRRKNGKAGHSYGLDKSYDLGDLBGLDPTDEYQON
WNTKSGSLRELRETLNGAVTRYVDNNPFCGPDGYPLDICKDFLAGRKMCTLSEQ
LDVIESKRGVYCCRDHEHEIAWTFERSDKSYEHOTPEIKSAKFTFKGCKFPVP
LNSKVKIOPRVESKKTTEGPMGRISVYPVAPSQECNWHLSLTMKNCDEVSQWTC
DFLCAECGCTENLIEGPTTCGYLPTNAVVMPCAPCDPBGPEHSHVADYHNHNS
IETLRKRGKRCFCGGFYAYCYNKAIRYVPRASADIGSGHTIGTGNVEDSL
LETLSRVENINTVGFHLNEEVAIIAGFSASTSAPIDTILSKDYSKFTIVESCGN
YKVTGKPKVGAWNIQOQRSVLTPLCGFPSSQAAGVIRSIPARTLDAANHSIPDLQRAA

VTILGISQSRLVDAMVYTDLLTNSVIIIMAYVTGGLVQOOTSOWLSNLLGTTVEKL
RPIFIEAKLISAGVEFLKDAWEILKFLITGVDFDI VKGOIQVASENLMKPCKCFIDV
NKALEMCIQDVITAGAKLSNLGEVFI AQSGKLYRQCI RKGEOLOLLMPLKAPKEVT
FLEGSHDVLITSEVVLKNGLEALETFVDSFTNGCAI VGTPTVCYNGMLLEIKDEKQ
Y CALSPGLLATNNVFLKGGAPI KGVTFGEDVWEVQGVKNRIT FELDERDYLKNE
KCSVYTESGTEVTFACVAVAVKTIQFVSDLLTNGMIDLDDEWSAFYFDPDDAGE
ENFSRMYCSFYPPDEEEDDAECHEERIDETCEHEYGTEDDYQGLPFLFGASAEVTR
VEEEDDLDWDTTQSELEPEPEPTPEBPVNOQFTGLYKLTQNVLAICVDI VKEAOSA
NPWVIVNAANIHLKHGGVAGALNKATNGAMQKESDDYIKLANGPLTVGGSCLLSGHNL
AKKLVAVGNLNAAGSDIOLLKAAVFNFSQDILLAPLISAGI FGAKPLQSQVCQVOT
VRTOVYVNDLKVQVMDYLDNLKPRVAPKQEPENPTNDSKTEEXSVQKPDV
KPKAKACIDVETTLTETKFLTNKLLLFADINGKLKHDSONMLRGDMSFLEKDAAPYM
VGDVITSGDITCVIIPSKKAGGTETMLSRALKKVPVDEYITTPGGCGAGYTLSEAKT
ALKKCSAFYVLSEAPNAKEBILGTVSNLREMLAHABETRLKMPICMDVRAIMATI
ORRYKGIKTQEGIVDGVRFPPFTYTSKPEVASIITKLSNLEPLVTMPICMDVRAIMATI
EAKRMSRLKAPAVSVSPDATTYNGYLTSSKTSSEHFEVTVSLAGSYRDSYSG
ORTELGVFLKRGDKIVHTLESPEFHLDGEVLSLDKLSLSLSEVKTIVKPTTVD
NTNLHTQVDMSTMKYQQGFTYLDGADVTKIKPHVNHGKTFVFLPSDDTLRSEAFE
YHTLDESFLGRYMSALNHTKWKFPQVGLTSIKWADNNCYLSVLLALQQLSEAFN
APALQAYYRARGDAANFCALILAYSNKTVELGVDRETMTLHQLHANLESKRVLN
VVKHCGQKTTTLTGVEAVMYMGTLSDYDLNKTGVSIPCVCGRDATQYLVQOESSFVMW
VFTKETS YTTIKPVS KLDGVYTIETIEPKLDGYKKNAYTEQPIDIVPTQPLNA
SFDNFKLCTSNTKFADNLQMTGFTKPSARELSVTFPDLNGDVDAIDYRHYSASF
GAKLHKPTVHINQATTTKTPNTWCLRMSTKPDVTSNSFEVLAYEDTQGMNL
ACESQPTSEEVNPTIOKEVIECDVKTTEVGVNVLKPSDEGVKVTQELGHEDLMA
AYMLFTKFFYLLGLSALIMQVFFGFASHFISNSLWMFIIISIVQAPVAVMYRIFF
AGFYIWKSVYHIMDCSTCMWCYKRNATRECTIVNGMKRSFYVANGRGFC
KTHNNCLNCDTCTGCTSTFIDEVARDLSLOKRPINPTDQSSYIVDSVAVKNGALHL
YDKAGQKTYERHPLSHFVNLNLRANTKGSPLPINIVEDCKSKDEASASVYV
SQMCOPIILLDQVSDVDSTEVSRMFDATVFTTSATFVPMKHLKALVATAHSE
LAKGVALDGLSTFVSAARQGVVDVTKDVI ECLKSHHSDLEVTGDSNCFMITY
NKEVNMPTRDGLACIDCNARHINAQVAKSHNSL IWNVDYMSLSERQKRSAAK
NNIPRLTCTATQVNVNITKISLGGKIVSTCKMLKATLLCVLAALVCYIVMPV
HTLSHDGYTNEIIGVVAIQDGVTRDIISTDDCFANKHAGFAMFQSFQSKYNDKSC
PVVAALITREIGTIVPGTGLVRAINGDFLHLPVFSAGVNICVTPSKLI EYSDFA
TSACLVAAECTIFKAMGKPVYCDTNLLEGSISYSELRPOTRYVLMGDSIIQFPNT
YLGSVRVVTTFDAYCRHGTCERSEVGLISTSGRWLNNEHYRALSQVFCGDMWN
LIANTPTPLVQPGALDVSASVAGGIITAILTVCAAYIPKFRFVEYNGVFNANAL
FLMPTTILCLPAYSFPLGVYSVFLYITFYFTNDVSPFLAHLQFAMFSPVFPVIT
AIVFCISLKHCHWPFNNYLRKRVMPNGTFTFEEAALCTFLANKEMYLKRSLELL
PLTOYNRVYALNKYKFSGALDTSYREAACTHAKALANDFSNADVLVYOPPTSI
TSAVLOSGRKMAFFSGKVEGQVCTCTTLINGLWLDVTYCEPHVICTAEDMLNP
NYEDLLIRKNSHFLVQAGNVQLRVIGHSMQCLLKLKVDTSNPKTKPKTKFYIQGQ
TFSVLACVSGSVYQCAMRNHTIKGSLFNGSCSGVFNIDYDCVFCYMHMELP
TGVAHTDLEGKFGYFVDRQTAAGTDTTITLAWLAWYAEVINGDMQNGRTTIT
LNDFNLVAMNYVEPLTDHVDILGSLTAQGTIAVLDMAALKEILLQNGMRTILGS
TILEDEFTPDVROCSTVFOGKFKKI VKGTHWMLTFLTSLILLVOSTOWLSLFF
VYENAPLPTLIGIMAIACAMLLVKKHAPLCLFLLPSLATVAYFNMVAPASWVRI
MTWLEADTSLSGRLKDCMYASALVLLILMTARTYTDAAARVWTLNMVITLVYKV
YTGNDQDAISMMALVISVTSYSGVTTIMFLARAIVFVCEYIYPLLTPGKTLQCI
MLVYCFGLGCCYCGFLGCLNRYFRLTGLVDYLVSTQEFYMSQGLLPPKSLSDA
FKNLKILGIGGKPCIKVATVQSKMSDVKTCSVLLSVLQOLRVSSSKLMAQCVOLH
NDILAKOTTEAPEKMWLSVLLSMQGAVIDNRLCEMLDNRAITQAATASRPSLPS
YAAATQAQAYEQAVANGSEVVLKLSLVAKSEFDRDAAMQRLKEMADQANTQ
MYQAKSEDYKRAVTSAMQTMFLTMLRLKLDNDALNNIINNAEDGCVPLNIIIPUTAAK
LMVVPDYGTCTKNTCDGNTFYASALWEIQVVDADSKIVQSEINMDSNLAWPLI
VTALRANSVAKLQNNELSPVALRQMSCAAGTTQACTDDNALAYNNSKGRFVLALL
SHODLKWARFPKSDGTGYIYTELEPCRFVDTPKGVKYLFTIKGLNLRNGVLL
GSLAATVRLQAGNATVPANSTVLSAFADPAKAYDKYLAGSGQPI TNCVKMLCTH
TGTCQAITVTPENMDQESFGASCLYCRCHIDHPNPKGFCDLKGYIQLIPTCAND
PVGFTLRNTVTCVGMWKGYCSQDLRPLMQSADATFLNRKVCVGSARLTPCGTG
ETDYNVRAFDIYNEKVAFAKFLNCCRFQKEBEGNLLDSYFVKHHTMSYQHE
ETIYNVQCPAVAHDFEFKRVGDGMVPHISQRRLTKYTMADLYVLAHFDNSGNDT
LKEIILVYNCDDDYFNKKWDYDFVENPDILKAVANLGERVRSLLKTLVQFCAMDMA
HGLVGLTLDNLQNLGMDVDFGVQVAPCGVPIVDSYSLLMPLITLTRAALAEESHM
DADLAKPLIKWLLKDYFTFEERILCLFDRYFKYWDQTYHPNCINCLDRCLLHCANFV
LFTSVFPPTSPGLVRKI FVDGVFPVSTGYHFRGLGVVHNDQVNDHNSRSLSKFELVE
YAADPPIHAASGNLLLDKRETTTCSVAALTNNAVFOQTKPGENFKDIFYDAVSKFPKE

```
GSSVBLKHFFPAQDGNAAI SDYDYRYNLPMTCDIRQLLFVVEVVDKVFDCYDGCCTIN
ANQVIVNLDKASGPPNKGRKLIYDMSYEDDODLAFATYKRNVIPTITOMNLIKTA
ISAKRATVAGVSGICSTMTRNQFHQKLLKLSIAATRGATVIGTISKFYGQNMMLKTAV
YDMEVTHGLMGWDYFKCDRAMPNMLRIMASIVLAKKNTCNSHRFYRLANECQAQVL
SEMVCPCGSLVYKFGGTSAGDATTAYANSVFNICQAVTANNVALLSDGKNIAKQYVR
NQHRLYECLVRNRDVEHFVDFEYAVILRKHFSSMILSDDAVVCVNSNYAAQGLVASI
KFKAVILYQNNPMSKACKWETDLTKGHPHPCSQHMTMLVKQDGDVYLYLPYDPSRI
LGACQFDDIVKTDQIMLIERFVSLAIDAYPLTHPNOEYADVPHLILYIRKULHDEL
RPHLMDYVSMVLTNDNSTRKLYSRYWEPFEYAMPTHTVLOAVGACVLCNSQTSRACACIR
TFLCCCKCCKDWTSTGHLKSLVSNPYVYCNAPGCDVDTQVLYLGMGSYTKSKHPRI
SFLPCANGQFGLKTCVSGSDNVDNFNAIATCDWTNAGDYILANTCTERKLIIPAAET
LKATEETFKLSYGATREVLSDRELHLSWEVGRPEPLARNRYVFTGVYTKNSKVOI
GEYTFEKGIDYDVAVYRGTTYKLVNGDYFVLTSTHVMPLSAPLVPQEHVYRITGLY
PTLNSIDFESNVANYIKQGMQKYSTLQGPFGTKGSHFAIGLADYFYSARIVYACSH
AANDALCEKALKYLPIDKCSRIIPARAVECFDKPVNSNLEQVFCFTVNALPETTAD
IVVDEISMATNYDLSNVARLRAKRYVYIGDPAQLPAPRTLLTKGLEPEYFNSVCR
LMKTIQPMFLGTGRCRPAEIVDTVSALVYDNKLAHKDKSAQCFKMFYKGVITHDVS
SAINRPOIVGVREFTNPAWRKAVFTSPYNSONAVASKILGLPTQTVDSQSGSEYDY
VIFTQTETAHSCNVNRPNAVITRAKIGILCIMSDRDLYDKLOPSTLSLEIPRNVATIQ
AIEVQTLGKDCSKLITGLHPTQAPHLISVDIKFTEGLCVDIPGIPKDMYTRRLISMM
GFRNYYQNGYPNFIITREEAIRHVRAMIWGFVDEGCHATRDAVGTNPLQLQGFSTGVN
LVAVPTGYDVTENNTEFTRVNAKPPGQDFKHLIPLMYKGLPMNVVRIKIYQMLSDTL
KGLSDRVVFLWAGHGLTSMKYFKVIGPERTCCLDKRATCFSTSDTYVACWNHVSIG
PDYVYNPMDIVQWGTNLSQNHDOHCQVGHNAHVASCDAINTRCLAVHECFVKVR
DWSVEYPIIGDELVRNACRKYQHWVKSALLADKFPVLHIDGNPKAIKCVQPAEVEW
KFYDAQPCSKAYKIEELFYSIAHTHDKFTDGVCLFVNCVNDIRYPANAIKVRFDTRL
SNLNLPGCDGSLYVNRHAFTHPAPDSAPTNLQLPFFYYSIDSPCSHGKQVYSDID
VYPLKSAATCITCNLNGAGVCRHANEYRQYLDANMMISAGFSLWIKYQDFTYMLMT
FTRLQSLNVAYNVNGKHPDGHAGEAPVSIINNAVTKVDIGDVEIPEKNTULPVNV
AFELWAKENIKPVEIKLANNLGVDIANTVINDYKREAPAHUSTIGVCTMTDIKXP
TESACSSULTFDORGVQDVLFRNANGVLDITEGSVKGLTPSGKPAQASVNGVTILK
ESVKTQNYFKKVDIIOQLPETTYNQSRDLEDFKPSQMBETDFLELADBEFTYRGL
EGYAFEHIVYDGFSGHGLGLMIGLAKSQSDSLKLEDFIPMSLTVKQYFITYDAOT
GSSKCVCSVIDLDDLFVEIILKSODLSVISKVKVTDIYABISPLMKDGHVEFTFP
KLQASQAWQPCVAMENLYKQRMILEKCDLQYGENAVIPKGINMNAVKTOLCOYLN
TLTILAVNRYVIFHGAGSKGVAPGTAVLRQWLPCTGTLILVDSLNDPVSADOSTLIG
DCATVANKWDLIISMDYDPTKHTVKTENDSKGFFTYICGFIKQALGKSIAVKI
TESHWNDLQKLMGHGFWTAFVTVNVAASSEAFILGANLYGKPEQIDGYTHHANYI
FWRNTPIQLSSYSLFDMSPFKLGRGTAVMSLKENQINDIMYLSLEKGRLLIRENNR
VWSSDILVNN"
196..13344
/codon_start=1
/product="nonstructural polyprotein ppla"
/protein_id="AAQ01607.1"
/db_xref="GI:33285980"
/translation="MESLVGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
HLKNGTCGLVEKGLVLPQEPVFIKRSDALSTNKGHKVVELVAEMDGIYGRSGI
TLGVLPVHVGTEPTIAYNRLRNKNGKAGCHSYGIDLKSYDLGDELGTDPIDYEON
WNTYKGSALRELRLNGGAVTRYVDNNGCPGDPYPLDCIKDFLABAGSMCTLSQ
LDYTESKRGVYCCRDHEHETAMFTERSDKSYEHQTPPEIKSAKFDTPFGCEPKFVP
LNSKVYIQPRVEKKTGEGFIRSVYPVSPASPOECNNMHLSTLMKNCHDEVSWQTC
DFLKATCEHGTENLVIEGPTTCGYLPTNAVVRMPACODPEIGPESHVADYHNHNS
LETRLKGGTRCFGCVCFAVVCYKNRAYVPRASADIGSGHTGITGDNVETLNEDL
LEILSRVNNINIVGDTHLNEEVALIILASFASASTAFIDTKSLDYKSFPTIVBSCGN
IKYTKPKVPGAWNIGQORSVLTPCLGFPQAAAGVIRSIPARTDLAAHNSIPDLQRAA
VTILDGISEQSLRLVDMVYTSDDLNTSVIIMAVTGLQVQTSQMLNGLTGTVEKL
RPIPEWTEAKLSAGVELKDAWETLKELTGVFDIVKGOIQVADSNLKCVKCFIDVV
NKALEMCIQDVTIAGAKRLSINLGEVFIASQKGLYRQICIRKEQLOLLMPLKPKVET
FLBGSDHVTLTSEVVLKNGELEALPTVDSFTNGAIVGTPVCVNGMLLEIKDKEQ
YCALSPGLATNNVFRLLKGGAPIKGVTFGEDTWEVQGYKNVRITFELDERVDKLVNE
ENFSRMVCSFYPPDEEEDDAECEBEEIDETCEHEGYTDEDYQGLFEGFASAEYVR
VEEEDDLDLDTTEQSEIEPEPEPTPEEPVQFTGYLKLTDNVAIKCVDIVKEAQSA
NPMVIVNAAINHLKHGGVAGALNKAATNGAMOKESDDYIKLNGPLTVGGSLCSLGHNL
AKKCLHVGNPLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVOT
VRTVYIANDNKALYEQVMDYLDNLKPRVEAPQKEPPNPTEDSSEKSVVQKPDVY
KPKTKACIDETVTTLLETKTKMLLFLADNGKLYHDSQNMLRGEDMSFLKADPYM
VGDVITGDTICVYIPSKKAGTNTMLSRALKVYPVDYITITYPGOCAGYITPEEAKT
ALKKCGAFYLPSEAPNAKEEILGTVSNLREMLAHAEETRLKMPICMDVRAIMATI
ORKYKGIKIQSGIYDVYGRPFYTSKEPVASIITKLASNEPLTMDPIYVTHGFNLE
EAAKMSRLKAPAVSVSSPDATVYNGLYTSSSKTSEEHFVETVSLAGSYRDSYSG
COTRGLGVFLKMGDKIKIVHTLESVPEHLDGEVLDLKLKSLLSREKVTIKVFTVD
NTNLHTQLVDMSTYGOQFGFTYLDGADVTKIKPHVNHEGKTFVFLPDSDDLSEAFE
```

CDS

source

FEATURES

Location/Qualifiers

1..29573

/organism="SARS coronavirus Taiwan TC3"

/mol_type="genomic RNA"

/isolate="TC3"

/isolation_source="clinical specimen"

/db_xref="taxon:239758"

/country="Taiwan"

join(196..13329,13329..21416)

/note="ribosomal slippage"

/codon_start=1

/product="nonstructural polyprotein ppla"

/protein_id="AAP97879.1"

/db_xref="GI:33188325"

```
/translation="MESLVGVNEKTHVQLSLPVLQVRDVLVRGFGDSVEEALSEARE
HLKNGTCGLVEKGLVLPQEPVFIKRSDALSTNKGHKVVELVAEMDGIYGRSGI
TLGVLPVHVGTEPTIAYNRLRNKNGKAGCHSYGIDLKSYDLGDELGTDPIDYEON
WNTYKGSALRELRLNGGAVTRYVDNNGCPGDPYPLDCIKDFLABAGSMCTLSQ
LDYTESKRGVYCCRDHEHETAMFTERSDKSYEHQTPPEIKSAKFDTPFGCEPKFVP
LNSKVYIQPRVEKKTGEGFIRSVYPVSPASPOECNNMHLSTLMKNCHDEVSWQTC
DFLKATCEHGTENLVIEGPTTCGYLPTNAVVRMPACODPEIGPESHVADYHNHNS
LETRLKGGTRCFGCVCFAVVCYKNRAYVPRASADIGSGHTGITGDNVETLNEDL
LEILSRVNNINIVGDTHLNEEVALIILASFASASTAFIDTKSLDYKSFPTIVBSCGN
IKYTKPKVPGAWNIGQORSVLTPCLGFPQAAAGVIRSIPARTDLAAHNSIPDLQRAA
VTILDGISEQSLRLVDMVYTSDDLNTSVIIMAVTGLQVQTSQMLNGLTGTVEKL
RPIPEWTEAKLSAGVELKDAWETLKELTGVFDIVKGOIQVADSNLKCVKCFIDVV
NKALEMCIQDVTIAGAKRLSINLGEVFIASQKGLYRQICIRKEQLOLLMPLKPKVET
FLBGSDHVTLTSEVVLKNGELEALPTVDSFTNGAIVGTPVCVNGMLLEIKDKEQ
YCALSPGLATNNVFRLLKGGAPIKGVTFGEDTWEVQGYKNVRITFELDERVDKLVNE
ENFSRMVCSFYPPDEEEDDAECEBEEIDETCEHEGYTDEDYQGLFEGFASAEYVR
VEEEDDLDLDTTEQSEIEPEPEPTPEEPVQFTGYLKLTDNVAIKCVDIVKEAQSA
NPMVIVNAAINHLKHGGVAGALNKAATNGAMOKESDDYIKLNGPLTVGGSLCSLGHNL
AKKCLHVGNPLNAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVOT
```

VRQVYIAVNDKALIEQVMDYLDNLKPRVEAPQOEPPNTEDSDTSEKSVQKQVDP
KPKIKADEVTITTELLTKNKLFLADINKGLYDOSQNMLEDMSFLEKADPYM
VGDVYISGDTITTEVTSKKAQGTTEMLSRALKKQPVDBYITITYPEQOCAGVIT
ALKKCKSAFVLPSEAPNAKEIEPITVSNWIRLEMLAETREKMLPCMDVIRYATL
QRKYKQIOBGIUVGVFFVTSKEPVASITIKLNSLEMLTPMIPGVYTHGFMLE
EABCRWSELKAPAVSSPDADVATVYNGVILYSSKTSSEHFEVTVSLAGSVKSVSG
QRTELGVFLKAGDKIVYTHLSPVPEFHLGDVSELKDLKLSUREKVIKYPTTV
NTNLQLODMSMTYGOQPTVLOGADVITIKPHNVNHEGKTFVFLSPDDITLSEAFE
YPTLIDBSFJGRYMSALNTHKMKVPOQGGUITSIKWADNCTGSLVALLAQLEVRPN
APALQAEVYARADGANFCAKILAYSNKTVGELGDVNRNETHLLQHANLEASRVLN
VQCHGCOKQTTITLGEVAVMYGTLSDYDLNLTGVSIPCVCGRDADTVLAQOESSFVMV
SAPAEKYLQOGTFLCANETVGNVQCQHYTHITAKETLYRIDGAHLTMESEYKGPVTD
YVYKETSITITLIKPSVUKDQVYITELBPJQVYTKONAYITSQPIDLVPLOPLNA
SFNFKLTCNTFPAQDLNQMTFTKTPASRELSVFTPOLPDGNAVYDRHSYSAFKK
GAKLLHKPIVHNIHQATTKTPKTNCLRWSTKFPDVTGSDVSEVLAZDITQOMDNL
ACESQPTSEBVENLTIQKEVIEDCVTIEVGNVILKPSDEGVYTOELGHEDIMLA
AYVENTSITIKKPNELSLAGHTATIGIAETIAINSVPWSKILAVKPSLOQALWTSA
CARKLAQRVNNTPVYFTLLFQLCFTFKSTNSIRIASLPTTIIAKNSVSVAKLICDA
GINVYSPKESKULTAMLLLSLSDLSGLSGLTAPAEVGLVNSFGAPSVNGREVLN
NSNWTMPCEGSPCSTICLSDGLSDISYPAETITQVITISYKZLITILGAEWL
AYMLFTFVYLLGSLAMOVFFCYEKAPEISNSWLMFTIISVOMAPSVAMRYITFF
ASFYIWKSVLHMDGCTSCMWCYKRNARFVBECTIIVNGMKRSPVAVANGRGCF
KTHWNCNCDTCTGSTISLEIVARDLSLOKRPINPTDOSSYIVDSVA/KNGALHL
YDKAGQKTYLPLSHFVNDLNRANNKTKGSLPINVIFDKSGMDESKASVATIE
SQMCOPIITLLOQALVSDSTEVSVKNNKTSPLNVTFSATFSPKMEKALAVATISE
LAKGALVDGLSVTFAAARGQVVDVDTQVIECLKSHSDSDTGVSDVSEVLAZDITQ
NKVENMTPRDLGACINCARHNAOVCASHNVSLITVWQVYMSLEQURLKISAAKK
NIPPLRTCANTRQVNVNLTIKLSLGGKIVSTCFPKMLKATLCLVLAALVCYIMVPV
HTLSIHGVTNELLGYKALQDQVTRDITSDDCPANKHAGDAMFORSQSGYKNDKX
PVAAITIREIGFVGLGPVLAIRGADVLPFRPVSAGNICYTPSKSLISYOFA
TSACVLAACETITFDAMGPEYCVDTNLLGESSYSELRDPTVMDGSIITOPNT
YLEGSVNVTTFIDKCRGTCEBSEGTCLSGTRGVNNHNRALSGVFCGVDANN
LIANIPTLPQVQALDASVAVGZIALITVYTAASFYKMFKRFQESVNVHVAANAL
LFLMSFTLCLPAPASFLPSVAVGLIYETVNTDASYFLAHLQEWAMFSPVPEWIT
AIVYFCLISLKHCHWPNNYIRKRVMPNGWVSTSEBAALCTFLKKNWYLKASETLL
PLTQNYRLALNYKYKFSGLDVTISREAAACHLAKALQFVNSGADVILPQPTSIS
TSAVLOSQFRMAFSPSGKGVQVTCITTLNGLHLADDTVYCPRHVITCAEDMLNP
NYEDLITKSNHSFLVQAGNVNRLVHSGMNCLLRLKGLDTSNPKTYKFFVIRQEQ
TFSLVACNGSPGVTVQACNPNHITKGSFLNCGSKGVGFINIDYCVSPCVMHMLP
TGVAHGTDLSEKPEYFVDQRTQAAGTQITITVNLWALVAAYVINGDRFLNRPFTT
LNDPLVANKNYEPLTQHVILLGPASQGTQIVLDMCAALKELLQNGNRTILGTS
TILDEFEPTPDVQVPSGTQOFGKFKVKGTHHMLLITSLISLIVSQTSQMSLFFF
VYENAFPTLFTGIMATAACMLLVKHIAFLCTFLPSSATVAVFNNVYMPASWARI
MTWLEADTSLSGRLKDCOMVASALILMTARTVDDAARVMTLMVITLVKVV
YVGNALDQALSMVALISVTSNYSQVTTIIFLARAVIYVCVYEPILFTIGNTLOCI
MLYCFGLGCCYFGLFCLLYNRFLYTLVQVILVQVPRYMSOGILPKKSIDA
FKNLIKLUGIKGKPKIYKAVQSKMDVACTSVYLLVQQLRVSESSKLUQACQLHL
NDILLAKOTFEAPKGVSSLLSVLKKLKNLNAKEFEDRNLMDRATQIASFESSFTV
YAAVATAQAEYQOANGVSSVLLKKLKNLNAKEFEDRNLMDRATQIASFESSFTV
MYQASDEKRAKVTAMOTMLTMLRKLNDALNININNAARDLNNISPLITTAAK
MYVQDGYGTNKTCDGNTFYVASALWEILOQVVDADSKIVOLSEINNDSINLAWPLI
MTALANSAVLQNNELSPVALRQMSCAAGTQITACTDNLALAYNSKGRFVILALL
SDHQLKWARFPKSGDGTGITIELSEPCRFVDTPKGRPKVLYITIKGLANLKNMVL
GSLAATVRLQOAGNEVPANSTVLSFCAVADPAKYADVILASQGPITNCVMKLTCT
PGTGTLANTVTPEANMQESFGAGSCCLYCRCHIDHPNKGFDLKGKYVOIPTTCAND
PVGFTLNTVPEANMQESFGAGSCCLYCRCHIDHPNKGFDLKGKYVOIPTTCAND
TSTDVYVAFEDIVNEKYAFKAPFLMTCRQEQEBEGNLLDSVYVKRHTWYNQYHE
ETIYNLQCPADVADHFFKPNFLMTCRQEQEBEGNLLDSVYVKRHTWYNQYHE
LKEILYTNQCDVADHFFKPNFLMTCRQEQEBEGNLLDSVYVKRHTWYNQYHE
GIVGVLNQLDQNGVYDFGFOVQAGKGVPIYVDSYVSLNMLPILTLTRALAESHM
DADLAKPLIKWDLITQDEERLCLFDPRYKFKVQDTPHNCINCLDRCILKASFNV
YADDPHMASGGLVARKIVDQVPGVSTGVRFLGVQVNOQVNLHSSRFSKELLAV
FASVTPMTRASGNLIDTCTCFESVAALTNNAVFTQKPGNFQVDFVAVSGSPKKE
GSSVELKHFFFAQDGNALSDIDYRNLPTMCDITQLLFFVEVVKFYDQDGCCLN
ANQVIVNLDKAGFPKSNKARLYYDMSYDDEADALFATKRNITPTTOMNLKYA
ISAKQVARTVAGVISCTNTRQFHQLKLSIATRTQVITVYKFGYGHMNLKTV
YSDVTEPHLMGDPKCDRAMPNMLINASIVLARAHNTCANSHRFLYLANEAOVL
SEMYMKGSLYKPGQTSGGDATALANSVFNICOATVNNALLSDGKNIDAKTVR
NLQRLYECLYRNEDVHEFVEYFYLARKHPSMILSDDAVCNYSNYAAQGLIASI
KNKAVLYQNNVSEBKAETWITDKGPEFLCOHMLVQKQDVIPLTPDPDSRI
LQFAGFVDDITKDTGTLMTIRELADITLPTGHPNOEQYADVLYIKRHLDEL
TGHMLDNYSMLTNDTSTRSEVPFEMVTPHTVQVACVLSNQSOTLSGACGIR
RPLCLCKCKSDHVISLHKLIVSNVPEYVAPGCDVDTQVLXGMSNCSCKHCKPI
RPLCLCKCKSDHVISLHKLIVSNVPEYVAPGCDVDTQVLXGMSNCSCKHCKPI

SFPLCANGQVPELKYKNTCVGSNDVDFNAIATCDTWNTAGDYILANTCTERLKLFAAE
 KRAETEETKLSTGIAVAREVLSDHNGVLSWEVKHPPRNLRNTFTGIRVTKNSKVLI
 GEYTFEKGDYGVVYRTTAKLVNDGYFVLTSTMPPLSAPTVLPQEHVYRTGL
 PTALISDEKAGVYKQVMQKYSTLQGPCTGKSHFAIGLALYPSVIRVATTCASH
 AADALCBKALKYLPIDKCSRILIPARARVECFDKFKNSTLEQVFTSVVATPEYV
 LVFDEIISWATYNDLSVNARLAKHYVYGDPAQLAPRTLLTKGTLEPEYFNSVCR
 LMKTIQPMFTGTCRCPAIVDTVSALVYDYNKLAKHDKSAQCQPMFYKGVITHDVS
 SAINRPQIGVREFITPRAMPKAVIIPSYNSQNSAVSKILGPTLPVDSQSGSEYDY
 IYFTQTTBESKNVNRNARIATRAKILGICTMSDRDLYDKLOFTSLETIPRNARVLI
 AENTVTLFGDCKSITILGHPTAQTHLSYDIKFTEGLCVDYIPGIPKDMTIRVLI
 GKMYQVNGYVNNPIITREEA LRHVRAVMPFDGVCYHARDAVGNVLPGLQSTGVS
 LFAVNYGVDTNENFTFRVNAKPPGQDFKHLIPLMYKAGLGNVVIRIKIOMLSDTL
 KGLSDRVVFLWAHGFEILTSMKYFKVILGPERTCCLDKRATCESTSDTYACNHSVG
 FDSVYNNPMDVQQMGFTGNLQSNHQBCHQVGNHNAHVASCDAMITCLRAVHECFVKRV
 DWYVYPIIDELRNARSKVQWNSWYKALLADKFLVLIHDGNPKAICKCVAPKRV
 KYFDAQPCDGSLYKYNKHAIFYTAPADKSAFTNLGOLFPFVYSDSPESHGKVVSDIL
 SNLMLPGCDGSLYKYNKHAIFYTAPADKSAFTNLGOLFPFVYSDSPESHGKVVSDIL
 VYPLKSAITCRNLGGVACHRIANEYROYLDANVMMISAGPSLWILYKOFDITNLANT
 FTRQLSENVAIVNNYKGVHFDGAGEAPVSIINNAVTKVADHVIPEIKNTTLPVNV
 AFELWAKNRIPKPEIKILNNGULDIANTVLIINWYKREAPHVSTIGCTMTDIAPK
 TESASCSNLTFLVDGVEGCVQDLFRNARNGVLTGEGSKVSTGKSPQASNVGVTILG
 ESKVTSFYNNFKKVDYIIOQLPEYTFQSRGLDEBFPGSRQMETDYGFLAMDEFTORYLK
 EGYAPFCHVITGVEGHQGLGHLMIIGLAKRSVDSPLVLTDAEISFMWLKQDHVETPE
 GSSKCVSVIDILLDDPVSIILKSQDLSISKVTVKIDVABEISFMWLKQDHVETPE
 KLQASQAQPGVAMPNLYMQRMMLLEKDLQNGENAVIPKGIIMNVAKYTLQCVLI
 TLTLVATNMRVWILISDMYDPRTRKTVKROSLPFTGLVDSLDLNFVSDADGSLILG
 DCAVHTKARNDLILISDMYDPRTRKTVKROSLPFTGLVDSLDLNFVSDADGSLILG
 TEHSNWDALYKLMGHFWSWTAFTVTVNWSASSBAFLIGANYLKPKREQIDGYTHMANYI
 FWRTNPTQLGSSYLSLFDMSKPKLKLKRVATVMSLENQINDMTIYSLLEKRLIITRENNR
 VVSSDDILVNN"
 196..13344
 /codon_start=1
 /product="nonstructural polyprotein pp1a"
 /protein_id="AAP977880.1"
 /db_xref="GI:33188326"
 /translation="MESILVIGNWEKTHVQLSLPVLQVRDLVGRGDSVEEALSEARE
 HLKNGTCGLEVEKGLPOLQPPVYIIRSDASLTHNGHKVLELVAEMDGIQVGRGSI
 WNTXGSLALRETPYAIRNLVNNKNGKAGCHSHYGDILKSYDLDELGDTPDIEDYEQN
 LDYIIESKRVYCKRDHEEIIAMFTSRSDSEYBQTFEIKSAKFTDPFGKECDKFPVP
 LNSKVVYQPRVEKKTGEMFRGIRSRVSPVASPQCENMLAKKNCHECDKFSWOTC
 DFLKATCBHGTENLVIIEGPTTCGYLPNTAVMVPACADPBEIGPEHSVADYHNHSI
 LETLRKGTGRTCPGCVGYAVYGVCKRAYVMPRASADIGSGHTITGDNVTLNDEL
 LEILSRERNVINGVDPHLNEVAIILASFASATSFDTIKSLDYKSKFTIVESGCI
 YKVTGKRPQKAWNIGQQRVSLTPCGPSPQAVGIRSLIPARTLDQMSLIDEMSVAIPLQRAA
 VILIDGSEQSRLVDAMVYTDLLTNVIMAYVTGIRVLTQMSLQSNLLGTFTVEKL
 NPIFEMIIDQVTSAGELKLDAMEILKFTIGVTDYIKVGQIQAASNDCKVKCFIDV
 NKALEMICDQVTSAGELKLDAMEILKFTIGVTDYIKVGQIQAASNDCKVKCFIDV
 FLEGDSDHVLTLSEVRLKNGLEALETPVDSFTNGAIYGVTPCVNGMLMLETIKDEQ
 YCALSPGLLATNNRPLKGAPIKGVITQGVSDTLVWEQGVKNVRIPELDERVQVLNLE
 KCSVYBSGTEVTEFAVAAVAVTKTQGVSDTLNMGIDIDEMSVAIPLFDJDAGE
 ENFSMRCSYPPDPEEEDAEPEEIIIDETCEHYGTEDDYQGLPLPFGASAETVR
 VESEEMEDDTLQBSIEPEPEEPBPNQFTGYLKLNTAMIKCVDIIVKEAQA
 NPMVITVAANTHLHGKSGVAGALATNGAMOKESDILYIKLNGPLTVGSGSLCHNL
 AKKLVHGNPLNAGEDIQLKAAVENFSDMILLAPILSAGYFGAKPIQSLQVCQVTV
 VRQVYIANDKALYEQVMYDNLNKRVEAPKQEEPNTEDSKTEESKVSQKPDV
 KPTICADIDEVITLTLEETLTKULLFADINKGLYHDSQNNMLRGDSMFLKDAQFIM
 GGDVITSGDITCVVILSKKAGGTEMLKRALKGVPEYDIYITPGGCGATYLEEAKT
 ALKCKASVYLPSPDGRVAFBEILGTSPNLSREMLAHABETRLKMIICMDVRAIMATI
 QRKYKGIKIQSIVDYGVRPFTYTSKPEPNIIRIITKLNSLDELATMPIGVYTHGPNLE
 EAACRMRSILKAPVUSVSSPDATVYNGVLTSSKTSSEBHFVETVSLAGSVYSSG
 ORTELGVFLRGDKIVIQHTLESVPEFHLGDGVSLDKLSLLSREVITIKVPTVDD
 NNTLQLVDSMTYQQGPPTYLOGADVTIKPHVNHGKGTFFVLPSDDTLKSEAFE
 YHITLDESFYRAGRASALNHTKKWFPQVGLTSIKWADNNSVLSVLLAQQLLEKFN
 APALQEYIYRAGRADAANFICALIYASNKTVEGELGVDRMTWTHLHQHLESKRVLI
 VCCKHQKQTLTGVEAVAMGTLVSYNLKTGYSIPCVGCRDIAQYLVQOQSSVMMV
 SAPRACYLKQGTFLICANEYTGNYOCHVTHITAKETLIRIDGAHTMSEKYGKPDV
 VPYKETSITTIKPVSYKLDGVYTEIIEPLKDYVKKONAYTEOPIDLVPTQPLNA
 SFDNFLKTCSTNPFADDLNQMTGTPKAPASRELSVTFPDLNGDVAIDRYHSASPKE
 GAKLHKPIVHNIQAQTKTTPKNTWCLRWLSTKPDVTSFVLAVDEYGMONL
 ACESQOQTSVEEVNPTIQXEIVDEKTTATTEVGNVILKPSDEGVPTELQELHMA
 AVENTSITIIKPNLESIALGKTIATHGIAIRANSVPMSKILAVKPELQGAQIITSN
 A

TFSLACNGSPGVTQCAMRPNHTIKGSPINGSCGSGVFNIDYDCVSCFYMHMELP
TGVAHGDTGKGVGPPVDROTAAQTDTITLNLVLAAYIAVINDGRWFLNFRFTT
LNDPNAVMKXNYEPLTDHVDILGPLSAGTITLNLVLAAYIAVINDGRWFLNFRFTT
TIIEDEFTPDVVRQSGVTFQGFKKIVGTHHMLLTITLSILLILQVQNSLFFF
VYENALPFTLGTMAACAMALVGHAFCLFLPLSPSATVAFYFNMVYMPASWVR1
MTLELADLSLSGLKDCVMYASLVHLLMTARTYDDAARVMTLMNVITLVYK
YGNALDQCTSMWALVTSATNSGVVITIMFLARAVFVCEVYPLLFTKSTLOCI
MLVYDCLGCCCCFGLRLLNRYERLTIGVYLVSTQBFYVMSGOLPPKSSIDA
FKNLKILGIGGKPCIKVATVQSMDSVKCTSVLLSLQOLVESSKLMAOCVOLH
NDILLADTEAFKEMVSLVLSLQGVADINRLCEEVDNRATLOARASEFSSLP
YAAVTAQEAQVANGSEVLLKRLKSLAVKBEFDRDAARQRLKEMADQMTQ
MYQARSDEKRAVTSAMQTMFLMLKLNDALNNIINNARDGCPFLNIPIITTAAR
LMVYDPYGTNTCDGNTFYASALWEIOQVNDADSKVQISEINMDSNPLAWPLI
VTRALRANSAPKLNELSPVALRMSCAAGTCTACTDDNALAYVYNSKGRFVLAAL
SHQDLKNARPPKSDGTGTYITILEPFCRVDTDPKPKMYLFIKGLNLRGVL
GSUATLVYDQGNATEVPANSTVLSFCAFADPAKAYKOLASGGQITNCVKMLCTH
TGQATITVTEANMGSGFGASCCLYRCHIDNHPKGFCDLKGKYQIIPTCGTD
PVGTITVTCVGMKGSGCCDOLREPLMSADASTFKRVCVGSARLPCTGTD
ETIYNLKDPCAVADHDFKFRVYDGMVPHISRLTKYTMADLVVALRHFGNCDT
LKELIYVNCDDDDYFNKDDYDFVENPDILRYANILGERVRSILKTYQFCAMRDA
GIVGVLTLDQNGWYDFGFVQVAPGCGVPIVDSYISLLMPILITLALAESHM
DADAKPLIKWDLKLYDTEERLCDFRDFYKQDTYHPNCLINCLDRCLIKHANCNV
LFSTVPPTSGFLVRKIFVDPGVFVYSTGTHREFLGVDHNDVNLHSLRSLFKELLV
YAADPAMHAGSGLLDRKTCFSAALTNVAFQVKGPNFNKDFYDAVSGKPFKE
GSSVELKHFFFAODGNAAISDYDYRYNLPTMCDIRQLLPVREVVDYFDCYDGCIN
ANQVILNNLDSAGFPKNGKARLYYDSNSYEDQDALPAYTKENVIPITLOMXYA
ISAKNARTVAGVSI CSTNTRQFHQGLKLSAATRGATVIGTSGFYGWNHMLKTV
YSVETPHLMGWDPKCDRAMPNLMIRASVLARKHNTCCNLSHRFVRLANESQVLR
SEMVCMGSLYVPGGTSGDATTAYANSVFNICQAVTANVALLSTDGNKADKYVR
NLQRLVYECYLRNRVDHFEVDFYLRKHFSMILSDDAVCYNSNTAAQGLVASI
KNFVALYONNVFMSEAKCWTDLTKGPHFCFSOHTMLVKGDDVYVLPYDPDSRI
LGAGCFVDDIVKDTGLMTERFVSLADAYPLTKHPNQEVADVFLHLYQIRKLDHEL
TGHMLDYSVMLNDNSTRYWEPEFTEAMTTPHVIQVAGCVLCSQTSYSLRCGACIR
RPLCLKCCYDHVISTSHKLVSNVYFVNCAPGCDVDTVQLYLGMGMSYSSKHKPPI
SEFLCANGQVGLYKNTCVGSDNVDNFATCDTNDAGDYILANTCTERLKLFAET
LKATEBTKLGYIATREVLSDRELHSEWKGKPRPLNRYVFTGRVTKNSKQVI
GEYTFEKGDGAVVYRTTYKLVNDYFVLTSHVPLSAPLVPQHYVYRTIGLY
PTLNIISDESPVANVQKGMOKYSTLOGPCGTSKHEAIGLALYPSARIVUYTASH
AVALDCEKALKYLPIDKCSRIIPARARVECFPKFVNSTLEQYVCTVNALPETTAD
IVFEDISPMATNYDLSVNARLRAHYVYIGDPAQLPAPRTLLTKGLTPEYFNSVCR
LMKTIIGPMFLGTCRCPAEIIVDTVALYDNDKLRKAKESQAQCFMYKGVITHDVS
SAINRPOIGVVRLEPNPARKAVFISPYNSQNAVASKILGLPTOTVDSOGSEYDY
VIPTOTETASCNVARNENVAITRAKIGILCINSDRDLYDKLQTSLEIPRRNVATLO
AENVTCGLKDCSKIITGLHPTOAPTHLSVDIKETKEGLCVDIPIGIKDMYRLJISM
GFKMNTQVNGVYNNFIREAIRHVRANIGDFVEGCHARDVATGNLPQLGRSTGVN
LVAVTGYVDENTENBFRVNAKPPDQPKHLIPLMYKGLPNVNRKIVQMLSDTL
KGLSDRVFVLMWAGHFLTSMTKYFVKIGPERTCCLDKRATCFSTSDTYACWNHVG
FDYVYVPMFIDVQWQFTGNLQSHDQCOVHGNHVASCDALMTRECLAVECFVRV
DWSVEYPIIGDELRVNSACKVQHVVKSAALLADKFPVLHDIGNPKAIKCVPAEVEW
KPYDAQCSKAYKISELFYSYATHDKFTDGVCLFNCNVDRYPANAVICRDETRVL
SNLPLFCDDGSLYVNHAFHTPAFKSAFTNLKQLPFFYISDPSCESHGKQVVSIDIL
YVPLKATCTCNLNGAVCRHANEYROYLDAYNMMSAGFSWLYKQDPDTYLNWNT
FTRLOLNVAYVNVKAGHEDFHAGEAPVSIINNAVYKVDGIDVIFESKTLTPVNV
AFELKARNIPVPEIKILNLDGVIDAANTVWDYKREAPAHVSTIGCTVMDIAKPP
TESACSLTVLPGRVEGVDLFERNARNGVLTIGSVKGLTPSKGPAQASVNGVTKLIG
ESVKTQPNFYKVDGIILOQLPEYFTQSDLEDKPKRSOMETDFELANDDEFQRYKL
EGTAFEHIVYGDPSHGLQGLHMLGAKRSQSLPLEKDFIEMDSTVRKNFYITDAQ
GSSKFRVSIIDLLDQFVELIKSODLSVTSKVVKVTIDYABISFMLWCKDHGVETFP
KLQASQWQGVAMPNLKYMRLMEKEDQNTQENAVIPKGMVNAVYKTOCQYLIN
TLTAVYVNNRVTHFAGSGDKGAPGTAVALRQWLPGLTLLVDSDLNDVSDADSTLIG
DCATVHTAKWDLIISDMYDPRTKHTKENDSKEGFTYLCGFKIKOKLALGSAVYKI
TEHSWADVLKMGHSPSWTAFVTVNASSSEAFLIGANYLKGKQIDQGYTHWANI
FWRNTPIQLSSVSLFDMSKFPLKLRGTAVNSLKENQINDMIYSLLEKGLLIRENR
VVVSSDILVNN
21413..25180
/codon_start=1
/product="spike"
/protein_id="AAB86788.1"
/db_xref="GI:40457434"
/translation="MFIFLLFLTLTSGSDLRCTTDDVQAPNYOHTHSMRGVYYPD
EIPRSDTLYLTQDLPLFPYGNVTFGHTINHTFGNVPVIFPKDGLFYFAATEKSNVVRGV
FGSTMNKSQSVIIINNSTNVVIRACNFELCDNPFPAVSKPMGTQTHMTI EDNAFNCT

CDS

FEYISDAFLDSEKSGFNKHLREFYKKNKGLFYVYKGYQIPDVRDLPSGFNTLKP
IFKPLGINITNFRAILTAFASPAQDIWGTSAAYFVGYLKPPTTLMKYDNGGNTITDAV
DCSONLAELKCSVKSFEIDGLYOTNFRVPSGDVREPNITNLCPGCEVFNATKF
PSYIAWERKISNCVADYSLVYNSTFFSTFKCVGSATKLNLCFSNVADSFVVKGD
DVRIQIAPGTVGIADYKLPDDDFMGCVLAWNTRNIDATSTGNYNRYKRYLRHKKRP
FERDINSNVPSDPGKCTPALNCPWPLNDYGYFTTGTIGYQYRVVVLSEFLLNAPA
TVCGKLSLDLKNQCNFNENGLTGTGLVTPSSKRFQPFQFGROVDFDTSVRDPK
TESELDISPCAFGGSVITPTGNASSEVALYQDVNCTNVSAIHADQLTPAWRIYST
GNVFTQAGCLIGAHEVDTSYECDIPIGAGICASYHTVYLSLLRSTSQSKSVATMSLG
AUSIYASNTIATPNTSISITTEVPVSMKTSVDCNNYICGDSSTECANLLUQGS
FCTQLNRLSGIAAEQDRNTRREVQVKMYTPTLKYFGFNFSOILPDLPTKRS
FIEDLNFVTLTADFMKQYCECLGIDINARDLICAKFNGLTVLPLPLTDDMIATY
AALVSGTAGWTFAGAAALQIPFAMQAYRFGICVGTQNVLYENQKAIANOQNKAI
QIOESLTTSTALGKLQDVVNQAOLNTLVKOLSSNFGAISVVLANDILSRLOKVAE
VQIDRLITRELQSLQTYVTTQIIRAEIRASANLAATKMSCEVILGSKRVDFCGKGYH
LMSFQOARHGCVFLHVTVPSPQRETTAPACHGKAYFPREGVYFNGTISWFIQ
RNFFSQIITDNTTNSGCDVVGIVINNVTVDPLQFELDSFGKGLDKYFNKHTSPDV
DLDGDISINASVNIQKEIDRLNEVARNLESILDLQELGKYEQYIKWPMYVWLGFIA
GLIAIWMVITLLCCMTSCCCLKGACSGSCCKFDDSDSEPLVKGKLVHT"
25189..26013
/codon_start=1
/product="unknown"
/protein_id="AAB82984.2"
/db_xref="GI:40557709"
/translation="MDLFMRFTLGSITAPQVVKIDNASPASTAHATATIPLOASLPFG
MLVIGVAFVAFQSATKIIALNKRWOLALYKGFQFCNLLLFVTVYSHLLLVAAAGME
AOFLYALYVLOACINACRIIMRCLWKCKSKNPLLYDANYFVCHWHTNYDVCIPY
NSVTDIVTEGDCISTPKLKDYOQGGYSEDRHSGVKDYVVVGHGFTVEYVQLSTQ
ITNTGIENTATFFFNKLVKDPNPVQIHTIDGSSGVANFAMDPIIDEPTTTTSPVL"
25610..26074
/codon_start=1
/product="unknown"
/protein_id="AAB86789.1"
/db_xref="GI:40457435"
/translation="MMPITLFAGTHITMTVTVHTVTSQQLSLLKVTAPQHQNSKKT
Query Match 90.9%; Score 20; DB 14; Length 29592;
Best Local Similarity 75.0%; Pred. No. 0.013;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GUGAACUCACUCUGAGCUC 20
|:||||:||||:||||:|
Db 697 GTGAACCTACTCGTGACTC 716
RESULT 16
AY395004
LOCUS
DEFINITION SARS coronavirus HZS2-Bb, partial genome.
ACCESSION AY395004
VERSION AY395004.1 GI:37624347
KEYWORDS
SOURCE
ORGANISM SARS coronavirus HZS2-Bb
SARS coronavirus HZS2-Bb
Viruses; seRNA positive-strand viruses, no DNA stage; Nidovirales;
Coronaviridae; Coronavirus.
REFERENCE
1 (bases 1 to 29620)
The SARS epidemiology consortium of Guangdong
From independent foci of epidemic outbreak to large genomic
alteration in late phase viruses: evolution of the SARS-coronavirus
Unpublished
2 (bases 1 to 29620)
The SARS epidemiology consortium of Guangdong
Direct Submission
Submitted (19-SEP-2003) Guangdong, China
JOURNAL
Location/Qualifiers
1..29620
/organism="SARS coronavirus HZS2-Bb"
/mol_type="genomic RNA"
/isolate="HZS2-Bb"
/db_xref="taxon:249089"

10. 11. 1991


```

RESULT 24
AY394996          29683 bp   RNA      linear      VRL 29-JAN-2004
LOCUS             SARS coronavirus ZS-B, complete genome.
DEFINITION        SARS coronavirus ZS-B, complete genome.
ACCESSION         AY394996
VERSION           AY394996.1 GI:37624339
KEYWORDS
SOURCE
ORGANISM          SARS coronavirus ZS-B
                  Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
                  Coronaviridae; Coronavirus.
REFERENCE
AUTHORS           1 (bases 1 to 29683)
CONSTRM           The SARS epidemiology consortium of Guangdong
TITLE             From independent foci of epidemic outbreak to large genomic
                  alteration in late phase viruses: evolution of the SARS-coronavirus
JOURNAL           Unpublished
REFERENCE         2 (bases 1 to 29683)
AUTHORS           The SARS epidemiology consortium of Guangdong
CONSTRM           Direct Submission
TITLE             Submitted (19-SEP-2003) Guangdong, China
JOURNAL           Location/Qualifiers
FEATURES          source
                  1..29683
                  /organism="SARS coronavirus ZS-B"
                  /mol_type="genomic RNA"
                  /isolate="ZS-B"
                  /db_xref="taxon:249081"

ORIGIN
Query Match      90.9%; Score 20; DB 14; Length 29683;
Best Local Similarity 75.0%; Pred. No. 0.013;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCGUGAGCUC 20
Db 761 GTGAACCTCACTCGTGAGCTC 780
      |:||||:||||:||||:|

RESULT 25
AY394997          29683 bp   RNA      linear      VRL 29-JAN-2004
LOCUS             SARS coronavirus ZS-A, complete genome.
DEFINITION        SARS coronavirus ZS-A, complete genome.
ACCESSION         AY394997
VERSION           AY394997.1 GI:37624340
KEYWORDS
SOURCE
ORGANISM          SARS coronavirus ZS-A
                  Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
                  Coronaviridae; Coronavirus.
REFERENCE
AUTHORS           1 (bases 1 to 29683)
CONSTRM           The SARS epidemiology consortium of Guangdong
TITLE             From independent foci of epidemic outbreak to large genomic
                  alteration in late phase viruses: evolution of the SARS-coronavirus
JOURNAL           Unpublished
REFERENCE         2 (bases 1 to 29683)
AUTHORS           The SARS epidemiology consortium of Guangdong
CONSTRM           Direct Submission
TITLE             Submitted (19-SEP-2003) Guangdong, China
JOURNAL           Location/Qualifiers
FEATURES          source
                  1..29683
                  /organism="SARS coronavirus ZS-A"
                  /mol_type="genomic RNA"
                  /isolate="ZS-A"
                  /db_xref="taxon:249082"

ORIGIN
Query Match      90.9%; Score 20; DB 14; Length 29683;
Best Local Similarity 75.0%; Pred. No. 0.013;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCGUGAGCUC 20
Db 761 GTGAACCTCACTCGTGAGCTC 780
      |:||||:||||:||||:|

RESULT 26
AY394983          29699 bp   RNA      linear      VRL 29-JAN-2004
LOCUS             SARS coronavirus HS22-A, complete genome.
DEFINITION        SARS coronavirus HS22-A, complete genome.
ACCESSION         AY394983
VERSION           AY394983.1 GI:37624326
KEYWORDS
SOURCE
ORGANISM          SARS coronavirus HS22-A
                  Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
                  Coronaviridae; Coronavirus.
REFERENCE
AUTHORS           1 (bases 1 to 29699)
CONSTRM           The SARS epidemiology consortium of Guangdong
TITLE             From independent foci of epidemic outbreak to large genomic
                  alteration in late phase viruses: evolution of the SARS-coronavirus
JOURNAL           Unpublished
REFERENCE         2 (bases 1 to 29699)
AUTHORS           The SARS epidemiology consortium of Guangdong
CONSTRM           Direct Submission
TITLE             Submitted (19-SEP-2003) Guangdong, China
JOURNAL           Location/Qualifiers
FEATURES          source
                  1..29699
                  /organism="SARS coronavirus HS22-A"
                  /mol_type="genomic RNA"
                  /isolate="HS22-A"
                  /db_xref="taxon:249068"

ORIGIN
Query Match      90.9%; Score 20; DB 14; Length 29699;
Best Local Similarity 75.0%; Pred. No. 0.013;
Matches 15; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GUGAACUCACUCGUGAGCUC 20
Db 724 GTGAACCTCACTCGTGAGCTC 743
      |:||||:||||:||||:|

RESULT 27
AY283795          29705 bp   RNA      linear      VRL 12-AUG-2003
LOCUS             SARS coronavirus Sin2677, complete genome.
DEFINITION        SARS coronavirus Sin2677, complete genome.
ACCESSION         AY283795
VERSION           AY283795.1 GI:30468043
KEYWORDS
SOURCE
ORGANISM          SARS coronavirus Sin2677
                  Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
                  Coronaviridae; Coronavirus.
REFERENCE
AUTHORS           1 (bases 1 to 29705)
                  Ruan, Y., Wei, C.L., Ling, A.E., Vega, V.B., Thoreau, H., Se Thoe, S.Y.,
                  Chia, J.-M., Ng, P., Chiu, K.P., Lim, L., Zhang, T., Chan, K.P.,
                  Oon, L.E.L., Ng, M.L., Leo, S.Y., Ng, L.F.P., Ren, E.C., Stanton, L.W.,
                  Long, P.M. and Liu, E.T.
                  Comparative full-length genome sequence analysis of 14 SARS
                  coronavirus isolates and common mutations associated with putative
                  origins of infection
                  Lancet 361 (9371), 1779-1785 (2003)
JOURNAL           Lancet 361 (9371), 1779-1785 (2003)
MEDLINE           22667074
PUBMED            12781537
REMARK            Erratum in: Lancet. 2003 May 24;361(9371):1832.
REFERENCE         2 (bases 1 to 29705)
                  Wei, C.L., Thoreau, H., Chia, J.-M., Chiu, K.P., Ng, P., Lim, L.,
                  Lambert, E., Zhang, T., Wong, M., Se Thoe, S.Y., Chan, K.P. and Ruan, Y.
                  Direct Submission
                  Submitted (27-APR-2003) Genome Institute of Singapore, 1 Science
                  Park Road, The Capricorn #05-01, Singapore 117528, Singapore
JOURNAL
TITLE

```


Job time : 1700 secs

THIS PAGE IS BLANK